

T H E S I S.

---

"On the Upper Limit of the Growth of Intelligence  
as determined by Mental Testing."

Presented for the Degree of Ph.D.

by

ELIZA ROBERTSON COLLIE

M.A., B.Ed.

---

May 1925.

*Degree conferred 22<sup>nd</sup> July, 1925.*



# C O N T E N T S.

	<u>Vol.I</u>	<u>Page</u>
Brief Historical Review of Mental Testing.		
Individual		3
Group		48
Performance		53
Mental Ability		60
Theories of "Intelligence".		72
The Controversy.		
Terman's versus American Army Results.		89
Green.		91
In support of the 16 year old level.		102
Against Terman's views.		115
Theoretical Considerations of Growth.		127
Parallelism between Mental and Physical Growth.		138
Historical Bearings on the Question of Mental Arrest.		144
DOLL'S "Growth of Intelligence".		157
	<u>Vol.II</u>	
Investigation.		
<u>Preliminary Test.</u>		183
James Clark School		206
George Watson's Boys' College		219
Edinburgh Ladies' College		226
Adults.		235
Combined Results		250
<u>Second Test</u>		266
James Clark School		282
George Watson's Boys' College		296
Edinburgh Ladies' College		303
Adults.		314
Combined Results.		341
General Conclusions.		359



## BRIEF HISTORICAL REVIEW OF MENTAL TESTING.

---

Mental testing, as a branch of experimental psychology, came into prominence at the beginning of this century. It is true that the honour of first attempting to measure 'intellectual faculties' falls to Sir Francis Galton in whose laboratory many investigations were carried on with individual subjects. His work and the work of Cattell in America might be regarded as pioneer efforts in the field. The earlier investigators were content to limit themselves to the study of simple mental processes, by means of single tests, such as the tests of sensory discrimination which are reported in Galton's *Inquiries into Human Faculty*.<sup>(1)</sup> Performance in these tests was found to correlate highly with intelligence as otherwise estimated and was little affected by practice or experience. Similar work, on a very extensive scale, was carried on by Binet and others in France, as may be seen from the many contributions to the subject in the pre-1905 volumes of *L'Année Psychologique*. Nevertheless, the whole of the early experimentation was characterised by two distinct lacks, the lack of the study of complex mental processes and the lack of adequate statistical procedure in dealing with the test/

---

(1) Galton. *Inquiries into Human Faculty*. P.23.

test results. Accordingly, it was not strange that the whole method of mental testing fell into disrepute for a time until two influences brought about a revival in the present century. The reaction occurred in France and England and was due largely to the work of Binet in the former country and Spearman in the latter. The new investigations differed from the old both in purpose and in method. "Their purpose was no longer to gauge any faculties, such as apprehension, discrimination, etc., but instead to evaluate general ability".<sup>(1)</sup> The attention was turned from the employment of single disconnected tests of simple mental processes, to the development of groups of tests of more complex mental processes. These latter were found to correlate more closely with intelligence estimates than the tests employed earlier. At the same time more accurate measuring devices were introduced for treatment of the results obtaining from testing.

---

(1) Hart & Spearman. British Journal of Psychology  
Vol. 1912-13.

## INDIVIDUAL TESTS.

## BINET

The Binet-Simon Scale<sup>(1)</sup> for the measurement of intelligence was the outcome of some fifteen years strenuous labour on the part of its chief initiator. For a considerable period prior to its appearance, Binet had interested himself in the development and employment of individual tests for measuring mental ability. Many of his results are reported in the early volumes of *L'Année Psychologique*. The scale, thus, was not the inspiration of a happy moment; it was the culminating point or the logical result of his early researches and it marked a stage in the development of intelligence measurement taking the place of many subjective methods of estimating ability, particularly the ability of subnormal individuals. In 1905, in response to the expressed desire of a commission appointed by the Minister of Public Instruction, Binet set about devising a series of tests which were to aim at discriminating degrees of mental ability. The Committee desired to detect all cases of acute mental deficiency with a view to segregating children possessing the defect into schools better adapted for their needs than the ordinary schools were. The original purpose of the Binet tests was thus to diagnose mental defect. The tests which Binet employed were varied in nature/

---

(1) Binet: *L'Année Psychologique* 1908 or *The Development of Intelligence* by Kite.

nature, the principles underlying their selection being an attempt to gain a many-sided or comprehensive view of a child's mentality and an avoidance of everything that savoured of school information or knowledge. As far as possible the tests were to give equal advantage to all children, without regard to social position, intellectual opportunity or other varying environmental factors. Regarding the change in the point of view which Binet exhibited, Hart & Spearman write "Binet's tests were not selected by a rigorous analysis of the intellectual functions but have as miscellaneous a character as possible. Further, evaluation is never based on the result of a single test but on the collective result of a motley group of tests".

In 1905, Binet sought to make his method of detecting subnormality purely psychological and to determine the level of intelligence on the basis of the relation of intellectual capacity to age. In this way, using his first series of tests, supposed to cover the chief psychological processes or traits that go to make up intelligence, he was able to make a rough classification of children according to their native ability, and, in particular, he was able to detect those who were exceptionally or markedly defective. Illustrations of the variety of the tests employed may be had from noting the requirements at each part of the series. At age five, the normal child is expected to be able to perform a triple order/



order, to draw a square from copy, to repeat syllables, give his age, and so on. At age seven, he should recognise missing features in a face, add three pennies and three halfpennies, state the difference between concrete objects, etc. At age twelve, amongst other things he gives three words to a rhyme, rearranges mixed sentences, and describes and interprets pictures.

It soon became apparent that it would be advantageous to have the tests in the series grouped according to 'years' and that a standard of normal ability for each year could then be obtained which would facilitate the classification of subjects. From this idea the first actual scale of intelligence originated.

In 1908, Binet produced the first version of his Echelle Metrique. It included most of the tests of the 1905 series, with certain additions. The tests were allocated to 'mental age' groups corresponding to the different life ages from three to thirteen years inclusive. For a test to be deemed characteristic of a certain mental age group, Binet had decided that it must be passed by a majority of the children of the corresponding life age. The scale construction is, therefore, purely arbitrary in character, since the allocations were made as the result of practical experimenting of the tests with children of different life ages. As regards the marking of individual tests within/

within any group, Binet adopted the "pass or fail" method and made no allowance for partially correct responses. Further, Binet altogether ignored the problem of estimating the relative difficulty of the several tests within each age group. The scoring is arbitrary on account of this failure to weight the test elements according to their individual worth.

In the scale, the term 'mental age' is first introduced to express the individual's level of intellectual ability. Theoretically, a child of six years, who is normal, should pass all the tests up to and including those assigned to the mental age level six on the scale but should fail to pass any others. In practice, the point of break-down was seldom found to be exactly at the end of any ~~all~~<sup>one</sup> age group of tests, so that allowance had to be made for additional tests passed. The method adopted was to allow a credit of one fifth of a year for every additional test passed beyond the group in which all were passed. Provided five tests exist at each age group this method is fairly sound. The final mental age of a subject was ascertained by noting the level at which he passed all the tests and by adding one fifth of a year for each odd test passed above that level without regard to its position in any group. Binet's deductions from his experimental work were that an idiot did not exceed the two year mental level, nor an imbecile the eight year/



year mental level, and the upper limit for moronity he placed at the twelve year level.

The 1911 Scale was a modification of the previous compilation in the light of experience and criticisms. The advent of the Scale created great interest and considerable diversity of opinion existed among psychologists regarding its worth. Many applications of it were made and the comments which Binet heard together with his own observations induced him to undertake a revision. The scale was obviously characterised by weakness at both extremes. Binet was well aware of this defect. Certain of the tests included in the 1908 version were too scholastic in nature and were eliminated. Other tests were repressed and their place filled with new tests standardised as before. Some were redistributed in the age groups. Additional tests at the upper end of the scale served the purpose, though inadequately as Binet admitted, of testing higher levels of intelligence. These were designated "adult" tests and included drawing from imagination the cuts in folded paper, giving the differences between abstract terms. This final revision of the scale contained fifty-four tests distributed theoretically (though actually with some differences) so as to place five tests at each age between the three year level and adult level. The 1911 form of the scale is the last, and most investigators would admit the best/

best, though up to the time of his death Binet continued to work out improvements. The original belief that the tests on the whole measured native ability apart from the effects of schooling, seemed to be justified by the first experiments with the scale. But, concerning the influence of environment and social status, Binet was led, partly by reason of the experimental findings of Decroly and Degaud in Belgium, and partly from his own experience, to realise that his original expectations were wrong. That there existed in reality some relation between test performance and social status he could not deny, but that this invalidated the scale he would on no account admit.

The Binet-Simon Scale, in spite of its suggestiveness, lent itself to a variety of adverse criticisms. Ayres,<sup>(1)</sup> for example, makes note of the tendency to over-emphasise the linguistic ability and minimise the importance of performance. He complains that there are included tests which depend too much on recent environmental experience and tests which require that the child shall be able to read and write. Sheer memorising and puzzle tests are over-weighted and there is too much emphasis on the ability to define abstract terms. Similarly, Meumann<sup>(2)</sup> offers a series/

---

(1) Ayres: The Binet-Simon Measuring Scale for Intelligence: Some Criticisms and Suggestions.

(2) Meumann: Vorlesungen. Vol.II pp. 278-283.

series of criticisms as follows:

- (1) The single tests are not rightly graded according to their difficulty.
- (2) The tests determine quite different capabilities.
- (3) Tests of each kind are not sufficiently numerous.
- (4) Abnormal children with quite special defects, mental or moral, are not detected.
- (5) The tests are adapted to a definite milieu.
- (6) The aim of the testing is not clear.

Freeman<sup>(1)</sup> in comparing the Binet method with the Point Scale method of Yerkes, goes deeper. Among the criticisms which he offers are these:-

- (1) The assumption of the Binet scale that the mental development of all individuals proceeds by similar stages is false.
- (2) To say that a correlation between different functions is the same for all individuals at a given stage is also false.
- (3) It is a mistake to imagine that each stage of mental development corresponds in turn to a certain physical age and that there is a correlation between the different functions at different stages of development. So, the age-grouping method is wrong and the method of all or none scoring.
- (4) The age group method cannot make due allowance for different rates of development.
- (5) The same mental age may result in the case of individuals whose records are dissimilar.
- (6) The personal equation of the examiner exerts a varying influence.
- (7) The scale cannot diagnose moral imbecility, dementia and intellectual degeneration.

These are only samples of the criticisms offered.

There/

There exist many more. Some are valid and some are otherwise. Those dealing with what the scale does not attempt to do may be disregarded, those dealing with what it does badly must be dealt with. Subsequent investigators have effected alterations in the tests used, in their allocation to different parts of the scale and in the scoring devices. But the basic principle of the scale's construction, so long as it is adhered to, limits the revisions. It is testimony to the suggested value of the scale that it continues to exist in the face of all opposition. Doubtless, it will live till something better appears in its place.

#### GODDARD

---

In America, Goddard<sup>(1)</sup> produced the first general revision of the Binet Scale for the purpose of adapting it to use with American children. The revision was the outcome of an extensive application of the tests to ~~many~~ <sup>many</sup> children in the ordinary ~~inst~~ schools. Unlike most investigators, he manifested a preference for the 1908 form of the Binet Scale, though he admitted certain minor respects/

---

(1) Goddard: The Binet-Simon Scale.



respects in which the later revision showed superiority. In his own revision he adopted the method of locating the individual tests, a point which Binet had overlooked. Certain of Binet's tests he rejected - others he re-allocated and, in addition, he introduced tests of his own but as a general principle he aimed at "following Binet's order as closely as possible". The great need, he maintained, was to extend the scale, and to procure suitable tests for the higher age levels. Perhaps the only tests available were tests of experience and it might happen that these were the best tests. He writes "We may some day conclude that the boy or girl who has had an opportunity and who has not conformed to the canons of civilised society, is fundamentally defective in the qualities necessary for a useful citizen".

The work of Goddard was chiefly confined to the application of the scale to ~~young~~ ~~children~~ ~~so~~ ~~that~~ ~~it~~ ~~is~~ ~~easily~~ ~~understood~~ ~~that~~ ~~the~~ ~~standardisation~~ ~~of~~ ~~the~~ ~~revision~~ ~~might~~ ~~be~~ ~~more~~ ~~accurate~~ ~~at~~ ~~the~~ ~~lower~~ ~~levels~~. Recent experimenters have shown that this is the case. Doll<sup>(1)</sup> suggests that there/

---

(1) Doll: The Growth of Intelligence P.10.

there is evidence of some inaccuracy in the location of certain of the tests even at the lower levels, probably due to "selection" of the subjects employed. But, on the whole, the scale is regarded as being satisfactory for use with defective children and less suitable for testing normal children on account of its "efficiency level" being at ten years.

In placing subjects, Goddard found it convenient to adopt a classification of intelligence similar to that employed by Binet. Idiots were those whose mental age was approximately from one to two years, Imbeciles could be regarded as Low Grade, Middle Grade and High Grade, aged three to four, five, and six to seven years respectively. Morons had a mental age from eight to twelve years. Above this level normal children were to be found. In the years that have intervened between the appearance of Goddard's revision and the present time, he has remained one of the most staunch supporters of the Binet-Simon method. The problem of feeble-mindedness constitutes a particular attraction for Goddard, as his writings <sup>(1)</sup> reveal. He emphasises strongly the dangers which "incapacity" carries with it in regard to/

---

(1) Goddard: Feeble-mindedness: its causes and consequences. 1914.



to the individual's inability to render himself an efficient self-supporting citizen.<sup>(1)</sup> His attitude is many-sided. While recognising the importance of the Binet-Simon test methods in diagnosing mental defect, he, at the same time, utters a warning against a too rigid adherence to statistical results. "Their application to the human problem is fraught with many and serious liabilities to error".<sup>(2)</sup>

#### BOBERTAG and CHOTZEN.

---

The interest of the work of Bobertag and Chotzen lies in the fact that they were among the earliest experimenters with the Binet-Simon tests, and attempted to adapt the scale for use with German children. In 1911,+ the former reported<sup>(3)</sup> on his application of the tests in 1909, to 300 'Volkschule' children, between the ages of six and twelve years, as well as to Spielschule children aged five and six years, children from higher schools between six and twelve years and Hilfschule children aged eight to fourteen/

---

(1) Goddard: Human Efficiency & Levels of Intelligence. 1920.

(2) Goddard: Preface to Feeble-mindedness p.IX.

(3) Bobertag: Über Intelligenz prüfungen nach der methode von Binet and Simon.  
Zeitschrift für Angewandte Psychologie  
1911

fourteen years of all social classes. In testing with the Binet Scale Bobertag employed a somewhat free translation of the tests, a fact of interest, since, either on that account, or in spite of it, he was the most successful investigator as regards the results he obtained. Early experimentation led him to question, both the value of certain of the tests in themselves, and the location of certain tests in specific age groups. The lower tests were too easy and the upper tests too hard. Further, he objected that equality of tests was not manifested within age groups. In a later article<sup>(1)</sup> he demanded definiteness regarding the placing of the tests on the basis of a percentage passing them. Binet had been content with an approximation of definiteness (between 60% and 90%). He expresses his belief thus "Es scheint dass es sich empfiehlt, folgende Forderung aufzustellen: ein Test ist derjenigen A.S. (Altersstufe) zuzuweisen auf der er bei etwas 75% aller normalen Kinder gelingt" p.505.

Bobertag was one of the first to employ the scale for re-test purposes and his results prove its validity for the purpose. He points out the weakness of/

---

(1) Bobertag: Zeitschrift für Angewandte Psychologie. 1912. P.495 et seq.

of the mental age method of expressing growth as evidenced by the fact that the number of children deviating either above or below the level of their age, increased as their age increased, so that less significance ought to be attached to a deviation either plus or minus at the upper ages than at the lower. Binet used two years' backwardness as an indication of real mental defect; Bobertag saw that the figures must vary. On this account, the I.Q. method of expressing feeble-mindedness (i.e. ratio of mental age to life age) appeared more just. His insistence on its worth, however, did not extend to a belief in its constancy, partly on account of the work of Chotzen<sup>(1)</sup> which indicated a decrease in I.Q. from age to age with children of the Hilfsschule who were more than three years mentally retarded. Bobertag's disbelief in the constancy of the I.Q. led him to think that mental defect was not absolute from year to year but increased with increase in age. This made him suggest a modification of the (then) recognised form of the curved growth for various levels of intelligence. Whereas diverging lines of growth had represented the progress of the different grades of

---

(1) Chotzen: Zeitschrift für Angewandte Psychologie 1912. P.411 et seq.

of intelligence, he substituted curves which indicated a gradual decrease in the I.Q. from point to point.

The chief interest of Chotzen's work lies in the comparisons which he made between normal and defective test scores, comparisons which led him to believe that the Binet tests had a varied value in bringing differences in intelligence to light, and that those tests which were best for discriminating the intelligent, tended to be the same tests as marked differences in social status amongst the children tested by the scale. As a result of testing and comparison of his own with Bobertag's results he finds that the order of difficulty of tests on the scale, for normal children and feeble-minded, does not always correspond. (The experimental work of Binet in England with the Binet tests supports the conclusion). Accordingly, his order of the tests as shown in his own revision of the scale for the earlier years, does not agree entirely with the arrangement which Bobertag finds most satisfactory.

STERN/

# STERN

---

Stern, whose name is most frequently connected with what he himself calls "Differentielle Psychologie" is none the less familiar in the realms of mental testing. His definition of general intelligence as "a general capacity of the individual, consciously to adjust his thinking to new requirements" has perhaps had more popularity with psychologists than any other definition. In his early contributions to the subject<sup>(1)</sup> he reviews the situation up to the time of writing, evincing an interest in the future of the Binet scale and its practical value for the practical problems of education. In criticism of the scale Stern finds an immediate point of attack in the tests themselves. According to the principle underlying the construction of the scale, the tests at each age should correspond to the mental phenomena which have just appeared at that age. That they do, cannot be maintained. The tests employed ought to bring to light differences which are otherwise known to exist. That this holds is again questionable. Further, there are too few tests at the later levels. The plus and minus method of scoring/

---

(1) Stern: Psychological Methods of Testing Intelligence. trans. by Whipple (1914)



scoring tests appears to be unsuitable. In the matter of allocating tests to the different levels he believes that it is necessary to discover "with what frequency the tests are passed at other age levels". The most satisfactory tests will be those which show a most decided advance with years. Again, in scale construction, the age levels must be so adjusted, that as many children test at age <sup>as</sup> and the rest fall equally on either side of the normal limits. The age levels, themselves, afford a grading that is too rough and fractional methods of rating ought to be introduced. In the second version of the scale Binet did not eliminate all the errors or inadequacies of the first.

Stern is not alone in his criticism of the scale's construction or in the matter of the tests employed. Where he is more original is with regard to his recognition of the inadequacy of the mental age method of expressing retardation. Since a year of mental age means less at the higher than at the lower stages of development it follows that difference between mental age and life age would not always give a fair estimate of backwardness. Accordingly, it is better to employ mental ratio. This is the/



the I.Q. which Bobertag almost simultaneously and independently devised. Regarding the question of its constancy, Stern is not in agreement with Bobertag and Chotzen. In a critical article<sup>(1)</sup> he maintains that the I.Q. is constant and that the results of various investigators (Bobertag, Chotzen, Descondres) are invalidated by their mistaken choice of subjects or mistaken analysis of results obtained from testing. He points out, and rightly, that the question of constancy does not exist with regard to the 'slowing down period' of growth, but only with reference to the period in which growth is progressing at its normal rate for the individuals in question. The study of feebleminded children suggests that they develop at a slower rate than normal children, and Stern would also maintain that they reach earlier their final level of arrest. The article concludes with a comment on the prognostic of the measure if its constancy be proved.

#### KUHLMANN.

In 1913, Kuhlmann presented his first revision of the Binet-Simon scale. More recently (1922) he/

---

(1) Stern: Zeitschrift für Angewandte Psychologie  
1916 P.1-18.

he has published his "Handbook of Mental Tests" in which the whole revision is set forth. The changes made in the original scale are considerable. An attempt is made to remedy the marked tendency of the Binet Scale to measure too low at the one end and too high at the other. A large number of tests are eliminated by reason of their unsatisfactoriness. The scale is extended at both ends and from two years upwards eight tests appear at each age group. In addition, an effort is made to standardise the procedure in testing. In two respects, at least, Kuhlmann has successfully overcome difficulties in the original scale. A correct judgment of mental ability was previously prevented by the over-emphasis on verbal responses and by inaccuracy in the scoring methods employed.

Referring to the improvements in his revision Kuhlmann writes "With the original scale and previous revisions of it, a reliable mental age could not be obtained much above the mental age of 10 i.e., for older children and adults near the border line between the normal and mentally deficient". Now, he has 129 tests, which are scored in terms of the time taken to do the tests and the number of errors/

errors made. Some of the tests are especially praiseworthy on account of the fact that success in them depends on the ability to concentrate thought.

Experimental results lead Kuhlmann to believe that the "development of intelligence stops at fifteen years, for all degrees of feeble-mindedness except idiocy," and for this reason he would employ the figure fifteen as the denominator, in calculating the mental ratio of individuals.

In the early days of experimenting, Kuhlmann was amongst those who originally conceived of the idea of using the I.Q. His reasons for employing it were, firstly that "the rate of mental development is not a decreasing variable but is constant through a succession of years" and secondly, "that in feeble-mindedness there is simply a retarded rate of development whose ratio to the normal remains constant." The first of these reasons is in contradiction to his general belief that the rate of development decreases with age for all grades. In reality, Kuhlmann's statements on this matter seem to suggest an evasion of definiteness. He probably supports Bobertag and Chotzen and recognises the use of the I.Q. without supporting the view as to its constancy, but, if so, his writings are not so clear as they might be.

DE/

# DE SANCTIS.

---

The work of DE SANCTIS followed very closely on the work of BINET himself. This investigator believed that the great defect of the Binet Scale lay in, what might be termed, its poor applicability to idiots, imbeciles and demented. This was understandable. Successful application of the scale to these 'grades' would suggest that "degrees of mental defect in idiots or imbeciles corresponded to the degrees of intellectual development in ages of growth". In the opinion of DE SANCTIS,<sup>(1)</sup> great weight was to be attached to "the stimulus of environment" as affecting the development of intelligence. Development could not take place of itself, yet man was born with the tendency to develop. DE SANCTIS believes in a certain amount of parallelism between mental and physical development, but since the indications are not always apparent, it is impossible to establish a direct relationship between age and mental defect.

In preference to the Binet Scale, DE SANCTIS/

---

(1) DE SANCTIS: Journal of Educational Psychology. 1911.

De Sanctis employed a series of personally devised tests and graded in difficulty, whose purpose was rather to measure the degree of mental defect, than the amount of intelligence that a child possessed. These are of a practical nature, involving the use of balls, and being practical, they are well suited to the testing of mentally defective children who are handicapped severely in language tests.

Investigators who have made use of the De Sanctis tests, note that they present in reality a series graded in difficulty, admit their applicability to children of the lower grades of intelligence, and suggest that they might act as a substitute for the Binet scale in certain circumstances without hoping to replace it. Very little support is given to De Sanctis' own view that the tests were applicable even to the years of adolescence.

#### TREVES-SAFFIOTTI.

---

The Treves-Saffiotti<sup>(1)</sup> Scale is an Italian modification of the Binet-Simon scale. Several opinions including their own based on a slight use of the test, led the experimenters to apply the tests widely and to criticise their worth. They tested about/

---

(1) TREVES-SAFFIOTTI: L'Année Psychologique 1911.



about 5% of the total school population in Milan adopting the procedure of BINET except in instances where they deemed it necessary to modify his technique. From the mass of results obtained in 1910 they were able to construct an order of difficulty of the tests, based on the percentage of children passing the tests. As a result of their experiments the investigators were led to propose a new method of evaluating and grouping the tests. They accepted BINET'S tests but rejected his method of evaluation, because the early tests were too easy and the tests were not arranged in the exact order of difficulty. They observe that it is "not possible to establish characteristic groups of tests for each class independently of the influence of age". So, the groups of tests must be presented for each age - but "en fonction de la classe fréquentée." That is, the first change they suggest, is the grouping of tests by school class as well as by age. The second is the classification of children on a qualitative rather than on a quantitative basis. This second point recalls the original outlook of BINET whose comprehensive grasp of the whole problem enabled him to view both its quantitative and qualitative aspects/



aspects. This width of view is not so apparent today. The qualitative aspect has retreated and is generally only implicitly considered - the quantitative is largely to the fore.

According to SAFFIOTTI'S method there are three sets of tests, easy, medium and hard. The grouping of children on the basis of their success was also three-fold. Those who succeeded in the simple tests were called deboli (dull); those who succeeded in the medium tests were called medii (average) and those succeeding in the third group were called forti (able). Further testing with tests belonging to other levels than the one to which the child actually belongs is permitted and is desirable if time permits. By this method a more accurate classification is obtained, by designating children very dull, average dull, bright dull, and so on.

As the result of their early experimenting in Milan, TREVES and SAFFIOTTI found this classification superior to BINET'S method. Their results revealed differences due to social status similar to those found by DECROLY and DEGAUD in their investigations with the scale in Belgium. In criticism of the scale they find it necessary to "recuser l'idée/

l'idée fondamentale d'une mesure d'intelligence sans culture". Intelligence develops under the influence of environment, and for children the environment is principally the school. They claim further that all the tests reveal the influence of instruction.

#### ROSSOLIMO.

---

ROSSOLIMO<sup>(1)</sup> is responsible for a variation of the usual method of assessing the mental ability of an individual. In place of presenting a final score to represent a subject's worth, he proposes what is called the Profile Method, a method which has found place in investigations from time to time since its first appearance, as for example in the recent work of ABRAMSON<sup>(2)</sup> in Poland and in the newly devised schemes of JUNE DOWNEY for investigating temperamental qualities. Mental processes to the number nine were studied including attention, imagination etc., and numerical evaluation (from 0-10 points) made of each process. The results for each individual were then combined in the form of a graph giving a kind of 'curve' or 'profile' characteristic of the individual. The claim attached to the method is/

---

(1) ROSSOLIMO: L'Année Psychologique. 1920. (Review of Beryl Parker's work)

(2) ABRAMSON: Fonctions Mentales de l'Enfant. L'Année Psychologique 1922.

is that it reveals specific weakness or strength which is not disclosed in the total result. The amount of strength or weakness can be compared with an estimated average for each process studied. Where the purpose of an experiment is to make dissections of this nature, as was the case with ABRAMSON, the method has much to recommend it. Its real value lies in the graphical representation of performance that it makes possible. But, where the purpose is to evaluate general ability it is not so commendable on account of its very nature, also on account of the time required for its application. The application of the tests requiring some three hours makes the method too ponderous for general use. Recently, <sup>(1)</sup> ROSSOLIMO has suggested a short method based on the same general arrangement, which occupies only between ten and twenty minutes even for use with backward children. RABINOVITSCH, experimenting with the method, commends it while pointing out the superficiality which characterises it and all short methods. She suggests its merit as an introductory test. The only other general criticism of the method pertains to the tests selected by ROSSOLIMO. It is doubtful if his selection is as good as it might be and some psychological <sup>is</sup> comment even on the non-psychological nature of some of the tests employed.

---

(1) Zeitschrift für Angewandte Psychologie. 1918  
(Eine kurze Methode)

# TERMAN.

---

In an early investigation with Childs, Terman took exception to the fifth of a year credit allowed by BINET for passing a single isolated test. He proposed instead to allow that fraction of a year as credit which corresponded to the number of tests within a group, since all groups did not contain five questions. At this stage, the writers suggested the use of ten tests at each age, with systematic repetition of tests, and, in addition, they recommended the use of the completion test originally employed by Ebbinghaus' definition of words, fables and other tests. The tentative revision was tried out on 310 public school children in three different towns, but since it did not prove successful, the necessity for a new attempt was recognised. The testing of 982 children helped towards the new revision of the scale up to age fourteen. Beyond that age Terman tested High School students and business men, besides a number of emigrating unemployed, chiefly between 25 and 40 years and juvenile delinquents of 14 - 19 years. The upper ages on the new scale were standardised, largely on the basis of the results of 400 "adults".

The Stanford Revision of the BINET Scale for which Terman was responsible, was published in 1917.

It/

It offers the most radical alternative ~~to~~ of the original, apart from the Point Scale method. Previous work with the scale had been of a more or less fragmentary nature, most investigators merely altering the place of a test here and adding or taking away a test there according to their particular findings. The Stanford Revision makes its chief contribution in the standardisation of the scale itself. Terman rejected in toto the method employed by Binet, he <sup>(1)</sup> writes, "We had already become convinced, for reasons too involved for presentation here, that no satisfactory revision of the Binet Scale was possible on any theoretical consideration as to the percentage of passes which an individual test ought to show in a given year in order to be considered standard for that year". Instead, he chooses that "The <sup>vi</sup>grinding principle should be to secure an arrangement of the tests on a standard of scoring which would cause the median mental age of the unselected children of each age group to coincide with the median chronological age". Accordingly, he deliberately adjusted the scale until he had achieved the necessary correspondence. It was only after three/

---

(1) Terman: The Measurement of Intelligence. P. 53-54



three revisions of the scale that an average I.Q. of approximately 100 was obtained for each age group.

The Stanford Revision contains ninety tests, six for each age from three to ten, for age fourteen, for the adult group and for the superior adult group, and eight for age twelve. In addition there are sixteen alternative tests. The revision succeeded better than other revisions in regard to supplying tests for the higher levels. New tests, either original, or derived from the experience of other investigators were included at different parts of the scale. In some cases, Terman improved on Binet's tests by supplying better material to develop his ideas. More <sup>trials</sup> ~~treats~~ were allowed the child in answering problems, and as regards scoring, there was introduced in some measures the idea of "partial credits" for partial success in certain tests.

From the application of the tests to a large number of children, Terman found that intelligence is fairly regularly distributed throughout the different ages from five to fourteen years. Above the age of fourteen there is lack of symmetry due to the operation of selective forces which eliminate poorer intelligence from the school environment. Within each normal group a great range of intelligence is to be/

be found but as far as fourteen years there is no evidence of increased variability in groups.

In estimating the retardation or normal or superior ability of the child, Terman employed the Intelligence Quotient (I.Q.) method suggested by Bobertag and Kuhlmann. This idea provided him with a convenient method of classifying subjects. The individual who tested at age, one, for example having a mental age seven corresponding to a life age seven, would have a mentalization of 1. or an I.Q. of 100 (as expressed in percentage form). In practice, individuals seldom tested exactly at age, in part, owing to the rough units of measurement of the scale, in part, also, on account of their natural variability. At any age group, assuming a normal distribution, the average I.Q. would be 100, but normality was a term applied to a range of abilities rather than to a single point in the normal distribution of I.Q.s. In other words, the individual whose I.Q.s diverged from the median I.Q. (100) ten points on either side (i.e. 90 - 109) fell within the category "average". Below 90 various degrees of feeble mindedness from "sub-average" to "idiocy" were/

were classified in the same manner and from 110 upwards superior ability was again graded. The scheme is in general use today. Its rigidity constitutes its chief defect, since all degrees of intelligence shade into one another in reality. But for practical purposes its value is unquestionable provided the defect is kept in mind. The possibilities of a 75 I.Q. can never be compared with those of a 130 I.Q. and it is safe to assume that differences apparent at one time of testing will not in general disappear at any subsequent period. It is not enough only to know the I.Q. as Terman would admit. The mental age of the subject must also be taken into consideration. The knowledge of the two affords an excellent means of sizing up the ability of a subject.

For Terman, the I.Q. presents a more significant indice than for some other investigators. Terman believes as the result of his own investigations that the ratio of mental to chronological age is, on the average constant, at least up to a point near the place where mental growth ceases which Terman believes to be "somewhere during the adolescent period of life". In calculating the mental/

mental ratios of adults he uses sixteen as the basis. The constancy of the I.Q. which various writers dispute is a very important matter educationally but at present the decision lies with the future. Suffice to say that Terman's faith is unshaken.

(1)  
A study of sex differences <sup>in</sup> ~~the~~ results from the tests revealed that, in the different age and grade groups, the girls tended to be slightly ahead of the boys. Between 7 and 10 years the difference was so slight as to be negligible but from 10 to 13 years the superiority of the girls was marked. After 14 years the boys took the lead. Terman explains the later subsequence of the girls not as suggestive of lesser ability, but as being due to the fact that a more rapid promotion of the girls handicapped them in the grade distribution. Analysts of Stanford Revision of the Binet scale suggest that it exhibits a marked linguistic emphasis which if true, would explain the superiority of the girls. "A linguistic bias would favour a linguistic sex". In general, the sex differences revealed/

---

(1) Terman. The Stanford Revision of the Binet Simon Scale.

revealed by the use of the scale are noticeably slight. Such differences as have been found pertain mostly to the performance of single tests within the whole.

Differences in social status are also recognised by Terman and the relation of social status to intelligence discussed.<sup>(1)</sup> He considers the importance of recognising the effect of environment on the scale itself but refuses to admit that the effect impairs the worth of the scale.

Although the Terman or Stanford Revision of the Binet-Simon scale, is one of the most radical, it is still open to criticism on certain grounds. Terman experimented for the most part with normal subjects and this enabled him to bring about a more adequate standardisation of the test at the upper ages, and to extend the scale farther than previous investigators had done. At the same time, the tests for the older years are still by no means satisfactory and they seemed to be too hard for the ages at which they are placed. The/

---

(1) Terman. The Stanford Revision of the Binet Simon Scale.



The inadequacy of tests for the upper ages is one of the difficulties which confronts all painstaking investigators and constructors of scales. That adequate tests cannot be procured, seems to suggest either a difficulty of devising tests for the higher age groups or an early creation of mental growth. In further criticism of the tests it may be added that those involving American phraseology and the use of American code are unsuitable for use in English Schools. Also, the tests employed do not always correlate highly with intelligence. (Terman at a later date warns against assuming too much from a test's high correlations with intelligence). There are many other of a more practical nature that might have been included in preference to certain of the too plentiful verbal tests of doubtful worth. Perhaps the greatest inconvenience of the scale is the time required for its administration. Thoroughness in constructing the scale pays the penalty that its application requires at least twice as long as the Binet scale tests.

YERKES/

## YERKES.

---

### The Point Scale

---

The Point Scale consists of fifty questions, grouped in the form of twenty tests. The questions are derived, with one exception, from the Binet-Simon Scale. "The Analogies" tests originally used by BURTT is the one addition. In the selection of the test material preference was given to tests "applicable through a considerable number of years", such as memory span and free association. Most of Binet's information tests were rejected.

In place of the pass or fail method of scoring which Binet employed, Yerkes substituted the method of partial credits for the subdivision of the tests, and even within the subdivisions themselves the maximum available score is 100. The total score for each subject is derived from the summation of his partial scores. Then, from a table of norms the chronological age at which the average subject would obtain that particular score is noted and/

and the mental age derived.

The advantage of the scale lies in its flexibility permitting of readjustment and revision of norms when desired. It is brief enough to be applied in full to each individual tested. The child simply starts at the beginning and is tested to the point of failure. In testing, the plan adopted is the giving of all tests of a similar kind together. Another advantage lies in the rejection of many of Binet's unsatisfactory tests. But certain of the poorer tests are still retained and certain others are overweighted. Still further the tests are not equally drawn from the sections of the Binet Scale, but some age groups have contributed more than others. Moreover, as regards the equality of the points there exists much doubt. Yerkes admits the lack of scientific weighting while stating, "The principle which we propose to employ in further development and perfecting of the point scale system is that of weighting in accordance with the correlation of a particular measurement with general intelligence or rather with the point scale score". He claims that the chief merit of the scale/

scale lies in its research and diagnostic value and in the ease with which it lends itself to statistical treatment.

In reality he would propose, not one, but three scales for mental measurement, one for the very young, one for the school ages and one for adolescents and adults, on the grounds that no one scale would supply the needs of all subjects. The principles underlying their construction are always the same.

The application of the Point Scale shows that the influence of social status on the results is marked, that sex differences are slight as regards the whole scale, but significant in the case of certain of single tests that race differences as revealed by a comparison of Hebrew and Irish subjects are hardly evident and that school training influences the scores but little.

The Point Scale method is young but it is suggestive of future worth by reason of its scientific basis. In certain respects it is undeniably weaker than, say the Stanford Revision, but in it there lies greater scope for improvement. YERKES himself/

himself observes its demerits but shows great faith in its possibilities. He writes optimistically: "I believe that the method of graded tests, . . . . will steadily gain ground, and that the gradation as gradually perfected will lead to unceasingly accurate standardisation - that there will accumulate reliable and varied norms for the guidance of the inexperienced examiner and that shortly our scales for mental measurement will consist of independently graded and standardised tests which can be used either alone for the measurement of particular response or in such groups as need dictates."

BURT.

---

The most recent revision of the BINET SCALE, that has appeared, is the outcome of the researches of Dr. CYRIL BURT during a number of years which were undertaken to supply the needs of an adequate standardisation of the scale for application to English children. Previous investigations in England by MISS JOHNSON and WINCH proved inadequate for this purpose. Discarding all revisions whatever, whether English or foreign, except for/



for the purpose of reference, on account of the dangers to which divergence from the original is open, BURT started with a translation of BINET's own work and adhered to BINET's descriptions as far as possible, "The modifications introduced into the procedure are almost entirely confined to such as are inevitably involved in translating the tests from French into English". Tests from both the 1908 and the 1911 forms of the scale were included in order to make the results comparable with those of other investigators. The application of the scale to English school children revealed the necessity for a rearrangement of the tests. The order of difficulty did not correspond with that from BINET's experimentation and neither did the age assignments. Accordingly, a thorough investigation was made resulting in the reclassification of many items. In December 1921<sup>(1)</sup> BURT presented his rearranged scale with meticulous directions at all points regarding procedure in application. The second part of his book discusses the validity of the whole.

Apart/

---

(1) BURT: Mental and Scholastic Tests.

Apart from the redistribution of the Scale elements for use with English children BURT recognises three comprehensive problems requiring solution. These are: -

- (1) "The average to be expected at each age from English School children, both normal and abnormal, and more particularly where the border line is to be drawn between the two types."
- (2) "How accurately do the tests, both severally, and as a whole, measure the intelligence of normal and defective school children and discriminate between the two".
- (3) "What influence is exercised upon the performance, apart from age and intelligence, by various extraneous factors - sex, social status, educational opportunity and emotional and moral dispositions".

The solution of the problems, as far as it was achieved, resulted from the statistical treatment of results obtained from the examination of 2674 normal children in London schools, 729 children attending schools for the mentally deficient and 107 juvenile delinquents from remand homes, industrial schools, and so on.

In this vast undertaking, BURT sifted thoroughly/

thoroughly the matter of the stability of the tests by determining the percentages of the children (normal and defective separately) passing the tests at the different ages on the scale. He determined the order of difficulty for normal and defectives and inquired into the equality of the intervals between tests in each group and tests in their relation to adjacent tests in other groups. The result of this part of the investigation led him to state "As regards equivalence of unit the number of tests passed provides a measure quite as valid as the number of mental years attained" but that "as a scientific scale the graduation of the BINET-SIMON tests, is not and cannot by any means be made exact."

The distribution of intelligence, for normal and defective children of each life age up to fourteen were obtained from the average scores. The results when compared with previous distribution relating to educational attainments <sup>(1)</sup> led to the conclusion "Intelligence varies more widely than school capacity". In addition, was observed an apparent/

---

(1) BURT: Distribution of Mental Abilities.

apparent tendency for the range of individual variability, as expressed in terms of the standard deviation, to increase with age. "In intelligence, as measured by the present version of the BINET-SIMON scale, children in London Elementary Schools tend to vary about the average for their age by exactly one year at the age of eight and a half, and throughout the earlier, if not the late, half of their school career, by nearly one eighth of their age".

The mental ratio mode of expressing mental retardation, BURT finds satisfactory on account of its independence of age. In defectives there is a slight tendency for the mental ratio to decrease with age, but this conclusion requires to be verified by subsequent experimentation. The growth of mentally defective children is subject to so much individual variability that caution must be exercised in making conclusions regarding it. On the average their annual rate of growth is only one half of the normal rate and they reach early their limit of development.

As regards the "central problem" of the work/

work, namely, the line of demarcation between normal and defective children, BURT's suggestion is that it should be placed at the point of "cleavage" between the normal and defective distributions of intelligence as expressed in terms of the standard deviation. This point occurs somewhere about 2.8 S.D. which corresponds roughly to I.Q. 70. But, further suggested difficulties (see P. 165) lead BURT to go beyond this conclusion and to state two postulates.

- (1) "That mental deficiency must be treated as an administrative rather than as a psychological concept".
- (2) That the line of demarcation for school children must be enunciated separately from the line of demarcation for adults.

In connection with the former he states, "For immediate practical purposes the only satisfactory definition of mental deficiency is a percentage definition based on the amount of existing accommodation. If in the special schools of London there is accommodation for only 1.5 per cent, then to adopt a border line which, followed out consistently, would cut off nearly 2 per cent, is clearly/



clearly indefensible.

The border line for children is taken at 70 per cent of the chronological age and for adults it is fixed at mental age eight.

The third general problem demanded a consideration of sex differences, social status, the relation of intelligence to educational attainment and vice versa, and a study of the test results of juvenile delinquents.

With regard to the sex question, BURT declares that "the two sexes seem during their intellectual progress to be playing a sort of statistical leap-frog, now one up now the other, during their whole school course". Separate age norms for sexes he considers hardly necessary.

Children of superior social classes are found to succeed better with the scale than others, a fact which, BURT insists, as did BINET two decades ago, does not diminish the worth of the scale. The influence of both social status and sex is largely confined to single tests within the whole, rather than to the scale itself.

In/

In considering the relation of intelligence to school attainment BURT observes that there is a marked tendency for educational agencies to concern themselves with the improvement of the 'not too dull' and to neglect, and consequently cause disinterest among, the children of superior intelligence. On the influence of educational attainments upon intelligence his remarks are (P. 182) "With the BINET-SIMON scale a child's mental age is a measuring not only of the amount of intelligence with which he is congenitally endowed, (Not only of the plans of intelligence at which in the course of life and growth he has eventually arrived; it is also an index, largely though not perhaps mainly, of the mass of scholastic information and skill which, in virtue of attendance more or less regular by dint of instruction more or less effective, he has progressively accumulated in school."

The application of the tests to juvenile delinquents evokes a warning against designating them all "deficient". Exaggerated views of the part played in delinquency amongst children, prevail and contributory factors are ignored.

The diagnostic value of the scale concludes the/

the investigation. For the most part the value of the single tests as diagnostic of normal ability, is low. "In discriminating the child of the special school from the child of the ordinary school, the scale as a whole is tolerably successful; in grading the special school children amongst themselves it is almost as efficient; in grading the normal children amongst themselves it is less accurate than other scales that are now to hand; and in detecting super-normal ability it is altogether invalidated by the anomalies and the lacunae among the problems for the higher mental years". In the meantime it is better than no scale and will lead to something more satisfactory.

The great value of BURT's work lies in the standardisation of the tests for use with English children and in the avenues his analysis has opened up for future investigators. It is only when a frank, clear exposition of the demerits of the BINET Scale is laid open for general perusal that a real understanding of its merits can be obtained. BURT has presented this exposition in "Mental and Scholastic/

Scholastic Tests." The imperfections of the measuring scale are plainly revealed but "when new buildings have not been finished, it is wiser to refrain than to demolish the old". Perfection in mental measurement lies with the future.

#### GROUP TESTING.

---

So far, the review of the historical position has taken into consideration only individual methods of testing. It was not until the time of the entering of America into the Great War that any other method existed, though about that time Otis was beginning to experiment in a new line. With the idea of testing the recruits of the Army in order to allocate to each man his appropriate duty, came the opportunity for group testing. Since then the development of group methods has grown apace. The convenience of group testing, permitting of the application of test to large numbers of subjects simultaneously, is largely to be considered in accounting for its popularity. The question of its relative worth as compared with the individual methods of testing is a different matter. The limitations/

limitations of the group method are considerable. For testing very young children, it is wholly undesirable and is not frequently employed.

The models upon which all forms of group testing are more or less based were constructed by the committee appointed to construct an intelligence scale for a speedy testing of the men entering the army from all the different avenues of life. Two forms of test appeared as the result of the committee's efforts, the one being the form known as Army Alpha, linguistic or rather, verbal in character, the other, Army Beta, of a practical nature, recommended for the testing of foreign speaking men or illiterates as well as for use with defectives who failed to do credit to the other tests. Both forms were carefully constructed and rigidly standardized.

(1)

The Army Alpha scale of tests in its final form contained eight varieties of test, namely, directions tests, arithmetical problems, tests of/

---

(1) YOAKUM & YERKES: Mental Testing in the American Army.



of practical judgment, synonym and antonym tests, disarranged sentences, number series completion, analogies and general information tests. A time limit of five minutes at most was attached to each section of the whole and the number of tests therein was such that not more than 5% of all those tested would finish in the time fixed. Fifty minutes was required for the whole performance. Alternative forms of the test were made to prevent coaching and cheating. The scores obtainable from the test ranged from 0 - 212 points and on the basis of the results the men were classified in seven grades from - "very superior intelligence" at the one end to "very inferior" at the other. The grades were marked A.B.C,+C,C,- D D -. As was expected, officer and non-commissioned officer material rated high when the tests were tried out on regular units, privates tested average and men unsuited for any service tested very low..

The Army Beta test represented, though not strictly, the Alpha in pictorial form. It consisted of seven types of tests, a maze test, a cube analysis test, X - O test, digit symbols, number/

number checking, pictorial completion and geometric tests. (Where necessary an abbreviated form of the Stanford Revision of the BINET scale or the Point Scale were also in use). Over two million men were tested by the Army test and the outcome of the whole immense piece of "mental machinery" was to create in America a strong belief in the possibilities of mental tests. The belief was in some respects over-great and psychologists had to check the over optimistic views of the army official and others and to see that the mental test estimate was not the sole criterion for designating a man able or otherwise.

From the employment of group tests in the Army it was but a step to introduce them into the schools and colleges with the result that group testing developed and new tests appeared every year. At the present time, the method is largely employed in America but extensive use of group tests has been made in England and some excellent tests have been devised on this side of the Atlantic. Some of the group tests employed resemble the Army Test in that they are miscellaneous in character. Others concentrate/



concentrate on one particular type of "ability", if it may be so called. Illustrative of the former are The National Intelligence tests, derived from the Army tests, the Otis Group Test, on which the Army test was partly based. The Northumberland Test of Godfrey Thomson and the Crichton Test of BALLARD. The two last mentioned are English compilations. The latter, concentrating on a particular type of test and presenting it in the form of increasingly difficult questions, is more restricted in nature. But the claim is that tests such as BURT's Reasoning Test and BALLARD's Absurdities touch rock-bottom and give a good estimate of an individual's native mental ability.

Group tests may be of the time limit or non-time limit variety or they may be a combination of both. The time limit variety preponderate, but in the light of much criticism of the method, suggestive of superficial tests and an emotional disturbance of the child, there are apparent efforts to adopt it with caution. Until the problem of the relation of speed to accuracy is determined the worth of these tests will be more or less assumed on account of the adequate correlation they usually yield/

yield with other estimates of intelligence.

A detailed study of the merits and demerits of group testing is hardly suitable here. Their great merit lies in the speed with which children can be tested and compared with the slow process of individual testing. If the purpose is to obtain norms of performance in any test or scale of tests their worth is unquestionable. But it is not to be expected that a short group test, and group tests usually are short, can give as accurate an estimate of each individual's mental ability as a more detailed, personal, oral examination. Much of the success of many group tests unquestionably depends on the adequacy of the directions which aim at making the situation clear to all those tested.

#### PERFORMANCE TESTS.

---

Up to ~~this~~ point reference to Performance Tests has been of the most casual nature. This type of test is, strictly speaking, of greater age than all others, going back at least to the middle of last century when SEQUIN produced his Form-Board for use with defectives. But, with regard to mental testing proper, this branch is decidedly backward/

backward. Practical tests find a place in the BINET scale itself and nearly all revisions of the scale have attempted to cater for the practical ability of the individual, but, so far their efforts have seldom been crowned with much success. Now and then a single test has met with some success as, for example, the Porteous maze test. It is composed of (1) 'mazes' of varying degrees of complication and the test requires that the subject start at the centre and with his pencil mark the way out. There are eleven mazes whose increasing difficulty corresponds to mental ages three to fourteen. The restricted nature of the material constitutes a drawback to the worth of the test which fails to yield as high correlation with intelligence as tests of a more verbal nature.

More success has met the efforts of KOHS, whose recently devised Block Design Tests yield higher correlations with the BINET Scale than other tests preceding it. The blocks used are sixteen in/

---

(1) PORTEOUS Maze Test: see BURT. Mental and Scholastic Tests.



in number, of one inch diameter and with all sides painted a different colour. Some sides are of a uniform colour, others have two colours diagonally placed. The material is said to make an immediate appeal to any subject. The test is to match designs with these blocks. There are twenty designs of graded difficulty drawn on cardboard. The modifications of the design are brought about by a number of devices. Each test has a time limit which is longer than the time demanded on the average. Scoring takes into account the time occupied in matching the design and the number of moves required. Penalties are given for over time and also if too many moves are made. KOHS claims that the test requires "a high analytic-synthetic ability of the mind" for successful performance. In spite of a high correlation of test results with BINET score ( $82 \pm .01$ ) KOHS declares, and with truth, "No performance test yet devised measures intelligence as adequately, and with as high a coefficient of efficiency as the BINET or others".

Scales of Performance Tests are few and far/

far between. Pintner and Paterson <sup>(1)</sup> have devised one to meet the needs of the psychological examining of deaf children. They also recognise the value of 'performance' as against 'verbal' tests for testing children from poor environments, where the language factor constitutes a problem. In construction the "scale" is loose. As they themselves say, "In a sense, the scale is a grouping together of tests in order to give a more general picture of the mental make-up of the individual." The attempt was made to obtain variety in the matter of tests and in this at least, the compilers have succeeded. General experimentation with the tests has revealed their inadequacy for testing the higher mental levels.

The same criticism applies to the tests of Healy and Fernald <sup>(2)</sup> constructed for psychological work in connection with the courts. Their value lies in their variety and in their applicability to/

---

(1) PINTNER and PATERSON: A Scale of Performance Tests.

(2) HEALY and FERMALD: Psychological Monographs Vol XIII.

to mental defectives.

There remains only the Army Beta Scale of which mention has already been made. Highly valuable as it was for the immediate purposes of the army, it has been subsequently found to be as inadequate as other performance scale and considerably less efficient than the parallel Alpha form.

The need for tests of a practical order is pressing but the difficulty of devising them constitutes a problem. There is the danger of making such tests too technical in character and the difficulty of choosing types that will make an equal appeal to both sexes. The suggestion that linguistic tests favour the weaker sex (physically) opens the way to a parallel suggestion that the performance tests would favour the boys. Whether or not this is true in regard to performance, interpreted generally, remains a matter of opinion. Certainly there are girls who can upholster chairs and boys who can sew on bottoms, and so on. In particular "performance" tests, of the nature usually devised, there is very little doubt that the boys are more successful on the whole than the girls.

The/

The main use to which performance tests have been put, is the diagnosing of mental defect. For this their success is unquestionable, since 'doing' is preferred to 'saying' by such children. LEROY STOCKTON, commenting on the inadequacy of performance tests for use with normal children indicates that their limitations occur at about the mental age limitation of defective children. He writes "When one looks at the amazingly uniform tendency of performance tests to reach the limit of their differentiating power at a point roughly shown in the graph (about 12 years) one must feel that it is probably more than coincidence that a mental age of 10 or 12 for adults has usually been chosen by intelligence experts as the dividing line between normals and feeble-mindedness. It seems possible that performance test standardisations have, perhaps unwittingly, established the approximate point when abstraction must gain the ascendancy or subnormally<sup>ity</sup> become apparent". There is some truth in the statement, but it seems to present a somewhat unfortunate point of view with regard to performance tests and to suggest a cleavage between them and other types of/

7

of test which does not, in fact, exist. The reason for quoting the passage, however, is that it reveals in part the inadequacy of the performance tests hitherto devised. Nearly all the PINTNER and PATERSON and HEALY and FERMAID Tests do indeed show a limit of improvability at about 12 years of age. But that this need characterise all performance tests devised in future is insupportable. KOHS Block Design Test has further possibilities. It recognises in fact, what LEROY STOCKTON ignores, - the possibility of different degrees of abstraction in performance.



MENTAL ABILITY

---

The recognition of individual mental differences was no new discovery at the end of the nineteenth century. In popular parlance terms such as idiot, imbecile and so on had their place, though doubtless the connotations attached to the terms varied considerably with the occasions on which they were used. What was then beginning to be recognised was the significance of these differences with regard to the competency or incompetency of individuals in practical life. It was only a glimmering of recognition which the procession of thirty years has served but to make slightly more clear and definite in the minds of the populace and the truth which wants yet to be thrust directly in the face of blundering Society.

From Galton came the first attempts to measure the degrees of ability which various individuals manifested, by the employment of definite individual tests credited to grade individuals according to their capacities. In the years that intervened between his initial interest in the question of mental testing and the appearance of the Binet-Simon Scale, it was more or less generally believed by investigators, that certain individual tests "tapped" the intelligence of subjects with a high degree of success. That the tests employed/

employed tended on the whole to be tests of simple mental processes was undoubtedly a disadvantage, as was later determined. Nevertheless, the first step was taken towards objectivity in measurement. The pioneer movements in the field of mental testing were characterised by a certain desultoriness of motive and lack of method which resulted in the scheme falling into disrepute until Binet recovered for it more than its original prestige. The foundationstones of the new 'Science' of mental testing were laid when the practical need for detecting decided degrees of subnormality resulted in the appearance of the first Binet series of tests in 1905. Single test methods were relegated to the background in the light of a new and more comprehensive conception of the use of mental measures. Series of tests were arranged at "mental age" levels corresponding to the different life ages and an individual was regarded as normal or otherwise according as he passed the tests of his own or other levels on the empirically standardised scale. Mental ability was now definitely recognised as constituting a part of the native endowment of man, as present in greater or less degrees in all individuals and as lending itself to measurement by the new test methods.

Further the idea of the development of this native ability presented itself through the same organisation of tests for the different mental levels or stages corresponding to the different life ages, and closely following on the new emphasis came the notion of/

of linking the terms, before applied vaguely to different types of intelligent being, to the definite mental levels or 'years' to which the different tests in the scale were allocated. The course of mental development of the normal child, from his earliest years to the stage of mental maturity could be observed from measurement taken from time to time and his normality would be approximately indicated by his ability to pass the tests corresponding to his particular age. Regarding the maturity of intelligence, Binet (1) writes "En quoi consiste cette maturité" Elle consiste en partie dans une augmentation de la faculté de comprendre et de juger..... elle consiste aussi dans l'augmentation du bagage des acquisitions de toute sortes. Mais se sont peut-être des caractères secondaires, et qui peuvent manquer sans compromettre la maturation" The absence of the quality of good judgment proclaimed defect which would show itself in the inability of the child to pass the tests of his corresponding life age. The term idiot came to be applied to those whose mental ability did not exceed the two years level, the term imbecile to those whose intelligence did not surpass the seven year level, and new terms "moron" (in French "débile" corresponds roughly) and "supernormal" were/

---

(1) Binet L'Année Psychologique 1908. P. 81.

were invented to designate respectively feeble-mindedness above the imbecile grade and ability beyond the grade into which normal individuals would naturally fall. Moreover, as a general working principle Binet took backwardness of two years in a child of any life age to be a clear indication of mental defect. The test method assumed a regularity in the development of native ability which subsequent testing has indeed proved to be true in the average but which in no wise applies to all individuals' mental growth. Particularly in cases of feeble-minded growth, as the work of Burt (1) reveals, are striking aberrations from the true course of growth shown. Binet (2) is careful to point the limitations of his own method. In regard to the mental age device for expressing backwardness he says "It has no bearing on the cause of the retardation, nor upon its particular nature, nor upon the means of rectifying it..... A defective child does not resemble in any way a normal one whose development has been retarded or arrested. He is inferior, not in degree, but in kind. The retardation of his development has not been uniform. Obstructed in one direction his development has progressed in others. In some extent he has cultivated substitutes for what is lacking; consequently such a child is not strictly comparable/

---

(1) Burt. Mental and Scholastic Tests P. 154.

(2) Binet. Mentally Defective Children p. 12



comparable to a normal child younger than himself".

Critics have voiced objections to the method of Binet, devised so as to yield a general picture of development. <sup>(1)</sup> These Seashore writes "Retardation does not follow a common flat level any more than growth does, nor even nearly so much. A child develops one capacity several times as fast and often at the expense of another faculty.. The differentiation is even more striking in retardation. What is more, those who employ the tests for practical purposes should not be satisfied with a flat mental age..... In a study of the normal individual we seek to discover his fortes and his faults, in short to discover his particular deviation from the norm of the common level. There is no reason why the Binet-Simon tests should not develop into specific measures of the relative rank, or age, of individual specific capacities and powers, such as reasoning ability, sensory observation, memory, imagination, initiative, emotional life, self-control, etc". Similarly, Pyle <sup>(2)</sup> argues "There is not in the Binet tests or any of the modifications yet proposed sufficient system - no common plan running through the tests for successive years..... We should plan a/

---

(1) Seashore: Journal of Educational Psychology  
Vol. III P. 50.

(2) Pyle: Ibid. P. 95.



a series of tests for determining the degree of development of logical memory, note memory, attention, imagination, association, and two aspects of mind more complex, learning capacity and reasoning..... It is more important, it seems to me, to know specifically the condition of the child with reference to the development of the separate mental traits than to know his average performance with respect to them all". Apart from their "faculty" savour, the truth of these assertions cannot be denied and Binet would have been the last to do so. But it is questionable how far they are radically opposed to the Binet-Simon methods and how far they diminish its worth. The only reply to such critics is that the method is still in its infancy and the belief is strong that something vital will emerge from the tentative scales now in use. Further considerations of the question of mental development are left over meanwhile.

Co-existing with the belief in the possibility of measuring intelligence and understanding its development, or better, existing as a radical part of it is the claim that the tests do not measure more than innate mental ability. It is asserted in favour of the mental test method that the knowledge which an individual has accumulated in the course of life, is left out of consideration and thereby/

66

thereby all individuals are placed on an equal footing. The mental Test is not designed to give an estimate of knowledge; it rates the individual apart from, and in spite of the actual knowledge he may possess. How far this claim is substantiated is another matter, regard to which investigators would probably be at variance with one another. Colvin (1) contributing his portion to the theoretical discussion on Intelligence in the Journal of Educational Psychology states "The tests in vogue are apt to be tests of Knowledge". Buckingham (2) writing in the same connection declares "By far the greater number of tests test Knowledge". These are extreme views but among serious investigators the influence of schooling or of environmental factors generally on test performance passes not unrecognised. As early as 1911, Goddard (3) realised the dependence of some of the Binet Tests on schooling and home environment but without pessimism. He notes "Perhaps there are no better ones (tests) than tests of experience and we may some day conclude that the boy or girl who has had an opportunity and has not conformed to the canons of civilised society, is fundamentally defective in the qualities necessary to a useful citizen".

---

(1) Colvin. Journal of Educational Psychology. 1921

(2) Buckingham Ibid.

(3) Goddard. The Binet-Simon Scale. 1911.

citizen". It must be noted that on the vexed question of opportunity for all Goddard has created a Utopia of his own, a society in which the most able minds will devise schemes and the less able help to put them into action! In more recent investigations Burt (1) has found it impossible to deny the influence of social status, of schooling and the play of environmental influences generally on the performance of children in the tests. The difficulty of determining the exact relation between powers that are innate and powers that are acquired through experience immediately presents itself. In this connection the view of Ballard (2) in the expression of his own opinion of intelligence merits recognition. He observes "It does not seem to me possible to maintain a clear cut distinction between powers that are innate and powers that are acquired and to say that intelligence is concerned with the former and not with the latter. For, the truth is; that every type of mental ability whatever its degree of generality or/

---

(1) Burt. Mental Scholastic Tests. p.

(2) Ballard. Group Tests of Intelligence.

68

or particularity is both innate and acquired, it is born with man as a potentiality, it is acquired as an actuality". The fact is that it requires stimulus from without or as Green would have put it "Intelligence does not exist in vacuo".

Ballard notes further that the distinction between innate and acquired **rests** on two factors, the first being "the facility with which the appropriate environment can arouse that ability to action, the second, "the degree of speciality in the environment itself". In short, his position is that Mental Tests cannot and "do not pretend to eliminate knowledge, but the knowledge is general. We cannot eliminate knowledge, we can only standardise its influence. Intelligence is the relative general efficiency of minds measured under similar conditions of knowledge, interest and habituation".

The insistence on the fact that the tests are not unaffected by external factors is the chief bent of the criticism. But, the criticism of Jones (3) throws another sidelight on limitations of the scale and its method of measuring intelligence./

---

(3) Jones. Some Problems of Adolescence.  
British Journal of Psychology 1923.



intelligence. "There yet remains much to be done in the way of critically examining the nature of many of these tests from the purely psychological side. They are, I think too often regarded from an external and matter of fact point of view, often from an exceedingly adult point of view, without due consideration being paid to the very subtle ways in which the reaction to the tests may be influenced by delicate, effective, and commonly quite unconscious factors. Speaking more generally still, I would say that my psycho-analytic work has given all a very high opinion of the intelligence of children and has **taught** me that there are some important respects in which it commonly becomes blunted with age or rather by the education which is an unescapable accompaniment of the increase".

Unfortunately, the control of unconscious factors is too much to expect from any method but the knowledge of their presence should constitute a warning against an over emphasis of the objective side of mental assessments.

All criticisms leave undisputed the sterling worth of the Binet Scale. Twenty years only have elapsed since its first appearance. Perhaps the wonder of it has dimmed in the process of time, /



time, or the promise of it became obscured in the complications that have arisen out of it, perhaps these have been side trackings from its suggested path, but it remains, to all who have the will to see it, a monumental discovery, pregnant with suggested <sup>use</sup> value for time to come. The **desirability** of adequately testing the native ability of children is great. One has not lived through a school life without being keenly aware both of the individual differences in the mental ability of children and of the dangers of subjective estimations of that ability. Dr. Burt believes that the best "estimate" of a child's intelligence may be had from a competent teacher who thoroughly knows her class. This may be true in general, the emphasis being upon the competency of the teacher. In particular cases, however, it is in question and it may be precisely in these cases that the need for a true estimate is greatest. When one considers from one's own experience how erroneous may be the estimate of a pupil's ability by a competent teacher, one is forced, with Burt and all leading psychologists, to pay due respect to the valiant attempt at "objectivity" which the advocates of mental tests are making. The claim of the new tests is this, and

and nothing more, that they can best detect the in-born mental capacity of the individual and can study with some accuracy its course of development by discounting the influence of schooling and all but the most general environmental factors. Testimony to their worth are the words of Yerkes: (1) "Mental tests are tools of some reliability in ascertaining the level of a person's ability to profit by experience and to adapt himself to changing demands. These tools are not 100% perfect, but they are reliable up to a certain point and will find a place in future needs".

---

(1) Yerkes; Mental Tests in the American Army.

## THEORIES of "INTELLIGENCE".

---

It is not proposed here to attempt a solution of one of the most enigmatical problems that confronts psychologists to-day. But, if there is one thing that must tire all who are interested in the problem of the nature of 'intelligence' it is to hear the familiar comparison of it with the problem of the nature of electricity. We have no definition of electricity, it is said, we do not know what it is, but we are content to believe that it can be measured, and, as some humorist has added, to pay for it when measured. So, of intelligence little is known and some are surprised that the world will not accept the fact of its existence without questionings as to its nature. The lack of a definition for each concept is as far as the comparison can be extended. What the populace is vitally concerned with is the practical utility of each. This established, definition is then a matter of no consequence. Writing a considerable number of years ago, Wells makes the following statement regarding the position of mental tests: "To justify themselves they must earn their bread in terms of usefulness for the questions of life". That is what is really demanded of the concept of intelligence and until the demand is satisfied the denunciator of tests must continue and those who employ them labour in the face of opposition. The root of the trouble lies in the/

---

(1.) Wells. Science 1913.

the fact that the opposition to mental tester, whether it be teachers or other educational agencies, is not clear as to what the concept implies, and small wonder at the circumstance, since so much confusion exists among psychologists themselves.

Strictly speaking, to effect an entire separation between the theoretical and practical considerations of a problem is impossible, but the desirability of concentrating on one or other of these aspects is, at times, so great as to constitute a danger that isolation may go too far. While it is true in the main, that the outstanding figures in the field of mental testing are those who have allied their practical work to their theoretical considerations of the problem of intelligence, it must also be admitted that a Tendency is evident for practice to run adrift from theory, from time to time, and vice versa, until some salutary influence brings them together again. The break away may be more apparent than real when a bulk of mental tests is hoisted on the market, it is charitable to suppose that the effect has no ~~mechanical~~ tendency but is the outcome of a real desire to **classify** the concept of intelligence itself and when theoretical discussions are carried on at great length it is presumably correct to infer that they have a direct bearing on practical methods.

(1)  
Galton/ In the early investigations, initiated by



Galton, the present day purpose of mental testing was absent. The object of mental tests as expressed in the familiar words of the pioneer, was "To obtain a general knowledge of the capacities of a man by sinking shafts, as it were, at a few critical points. In order to ascertain the best points for the purpose, the sets of measure should be compared with an independent estimate of a man's powers. We may then learn which of the measures are not instructive". By experimentation, certain simple sensory tests were found to correlate highly with intelligence and to be little affected by practice or experience. But, towards the general understanding of "intelligence" itself little progress was made, chiefly on account of the limitation of the tests employed and the lack of adequate statistical procedure, but also by reason of a distinct lack of aim in the testing practices. Even in this period of experimental drifting, theoretical sidelights on the problem of intelligence appeared in alliance with the practice of testing.

(1)

Ebbinghaus believes that "Mental ability demands not mere relative capacity, readiness of recall of facile association of specific past experiences, it demands all this and something more, something more complex and as it were creative, namely, the ability to combine into coherent and significant/

---

(1) Ebbinghaus. Zeitschrift für Psychologie XIII.

1. 401. Whipple: Manual of Mental and Physical Tests. P. 283.



significant wholes, mentally independent and even seemingly contradictory impressions. In short, intelligence is essentially a combinative activity. To measure intelligence, therefore we must employ a test that demands ability to combine fragments or isolated sections into a meaningful whole". Such a test may be afforded by the Kombinations-methode or completion test method which originated with Ebbinghaus. This test has the merit of being of the complex variety and the demerit of being but a simple test. The emphasis is on the ability to synthesise as Ebbinghaus' definitions "Intelligenz ist Kombinationsgabe" makes clear. A recent writer, KOHLS<sup>(1)</sup>, finds some kinship with this view, while pointing out its inadequacy and over emphasis of the synthetic activities of mind. He writes "Intelligence acts of all sorts require both an analysis of the situation which confronts us, a critical inquiry into a consistent whole. Our definition of intelligence as adaption (see Pater) fails to take into account the the higher synthesis and analyses of the thinker ... Intelligence more adequately defined would be expressed as the ability of an individual to analyse and synthesise. Ability in this direction varies as does physical height. Analysis and synthesis/

---

(1) KOHLS. Intelligence Measurement. Chapter V.

16

synthesis are the mechanics of intellectual behaviour adaptation to changing situations but are of the many results".

Theoretical considerations of "intelligence" were at the bottom of the revival of mental testing at the beginning of this century, though the real impetus was given to the movement by the appearance of the BINET series of tests. The change of outlook that characterised the new efforts is noted by HART and SPEARMAN<sup>(1)</sup>. "Their purpose is no longer to gauge any faculties such as apprehension, discrimination, etc. but instead to evaluate general ability". Immediately thus appeared a felt need to understand exactly the nature of the ability itself. Of the various hypotheses formulated, three are outstanding and it is convenient, at the outset, to consider the problem of the nature of intelligence from the three points of view so expressed.

Intelligence may be considered as a single intellectual function permeating all activities, or as a compound of two or three highly generalised faculties or functions, or as composed of a multiplicity of abilities all highly specific.

SPEARMAN<sup>(2)</sup> BURT<sup>(3)</sup> (and STERN) are generally quoted as upholders of the first, BINET probably/

- 
- (1). HART and SPEARMAN. British Journal of Psychology. Vol. 1912-13.  
(2). SPEARMAN. Ibid 1904.  
(3). BURT. All works.  
(4). STERN. Psychological Methods of Testing Intelligence.

probably fits in fairly well with the second, and Thorndike supports, with wavering fidelity, the third of these general views.

BINET shows what might be regarded as considerable reluctance to tie himself down to a definition of "intelligence", no doubt on account of his comprehensive grasp of the difficulties entailed. His later writings suggest that his view approximates to what is known as the "two factor hypothesis" namely the recognition of a general factor and specific factors separable from it. In his early writings (1) he states: "It seems to us that in intelligence there is a fundamental faculty, the alteration or the lack of which is of the utmost importance for practical life. This faculty is judgment, otherwise called good sense, practical sense, initiative, the faculty of adapting one's self to circumstances. To judge well, to comprehend well, to reason well, these are the essential activities of intelligence". So BINET accords the first place to judgment. In a later writing (2) analysing the 'thought process' or plan of operation of intelligence he observes 'Though, as we believe, is composed of three distinct elements: a direction, an adaptation and a criticism', and again (3) more specifically 'Comprehensive' invention, direction and/

- 
- (1). BINET. L'Annee Psychologique, 1908 - or The Development of Intelligence in Children - translated by Kite.  
(2). L'Annee Psychologique. L'Intelligence des Imbeciles, 1909  
(3). Les Idees Modernes sur les Enfants.

and the power of criticism (censura) intelligence lies in these four words". Further, - the specific mental progresses, as perception, discrimination and the like, are mere 'phenomena of intelligence' and controlled by intelligence. So much for Binet's approximations to definition. His profound analysis of the characteristics of normal and feeble-minded intelligence at the various stages of growth probably throws more light on the problem than all these put together, besides emphasising the qualitative aspect of intelligence which subsequent investigators have tended to efface more and more, especially in regard to normal children.

The method of procedure adopted by those who investigate the possibility of a single factor of 'general intelligence' permeating all mental activities, depends mainly on the employment of correlation. At the beginning of the century, in 1901, the method was first employed by Wissler in testing college students in America. He obtained an estimate of intelligence from the college marks of students in the various subjects which they studied and correlated this estimate with results from tests of simple processes. High correlations were indicated between the college marks in the different subjects of study and a failure to find significant correlations between the tests themselves and between test scores and college marks, suggested no common factor of a central character entering/



entering into all special capacities. Thorndike's work (1902), in which tests of processes closely allied to one another were employed, and which on that account might have been expected to yield high correlations, served but to confirm the findings of Wissler.

In 1904-05, Spearman, using few subjects instead of many as the others had done, and making use of a variety of tests of intelligence from the psychological laboratory, obtained contrary conclusions, and demonstrated mathematically the possibility of a General Factor. He obtained high correlations of specific abilities with general ability and inferred the presence of a common factor from the fact that he could arrange school subjects and tests in "hierarchichal" form, that is, so that the values diminished as the eye travelled from left to right, or from top to bottom of the page results. In a second investigation he found that the correlation coefficient between any two abilities increased with practice and so he concluded that the central factor was "a plastic function of the nervous system".

Following on Spearman's work came the early investigations of Burt <sup>(1)</sup> at Oxford. Using a series of/

---

(1) Burt. Experimental Tests of General Intelligence. (British Journal of Psychology Vol. III. P.P. 94- 117.  
Experimental Tests of Higher Mental Processes and their Relation to General Intelligence. (Journal of Experimental Psychology. Vol.I. P.P. 93-112.)



of motor, sensory and sensory-motor tests as well as tests of association and voluntary attention, and giving each test twice to obtain a measure of reliability, he obtained results which yielded correlations with other estimates of such a nature as to suggest that there was such a thing as 'general intelligence' and that it could be measured. The general factor was partly of the nature of a plastic function of the nervous system but its basis was also partly attentive. The early investigation suggested the desirability of developing tests of complex mental processes. In a later investigation at Liverpool, in which such tests were more generally employed, and most often group tests, the results of the previous experiment were confirmed. Intelligence, for Burt, then becomes "in-<sup>(1)</sup>born, all-round mental ability!" and "Of inborn general capacities intelligence is the most important. It can best be measured by tests requiring the voluntary maintenance of attention, quick and accurate learning<sup>(2)</sup> (adaptation) and on the higher levels".

The works of two additional experiments, of which mention is only made in passing, support the theories of a 'general factor', but each of the writers postulates the presence of a second factor also of a/

- 
- (1) Burt; Mental and Scholastic Tests. P. 199.  
 (2) Psychological Tests of Educable Capacity. (Report).

a general nature. Maxwell Garnett finds evidence of a factor which he would call 'cleverness' which, incidentally, Ballard scorns on account of its 'popular savour' and irrelevancy and Webb finds a new factor which he terms 'persistence of motives' and which he regards as 'permeating the character side'. Where the truth lies regarding the nature of intelligence, it is not possible to decide, but no doubt some portion of it resides in all hypotheses formulated.

Analogous to the view postulating a general factor is that of Stern <sup>(1)</sup> who complains of the width of the definitions of intelligence employed by psychologists which 'includes all the mental activities that are not volitional or emotional' and who suggests a further restriction of the term feasible only from the teleological point of view. "Intelligence" he maintains "is a general capacity of an individual consciously to adjust his thinking to new requirements: it is general adaptability to new problems and conditions of life". Stern's claim is that such a definition distinguishes intelligence from memory, from genius, "whose character is to create new spontaneously" and from talent "which is more or less specific in character". In the employment of it, it must be duly/

---

(1) Stern: Psychological Method of Testing Intelligence. (Whipple's translation 1914).

duly remembered that "only those phases of intelligence testing which deal with a scale of degree are involved". The importance of qualitative differences in types of intelligence should not be minimised. Despite Stern's attempt to restrict the meaning of the term, the change that is most frequently raised against the definition, is that it is too wide and inclusive of other aspects of mentality than intelligence pure and simple. Therein lies the difficulty of a teleological definition that it lends itself to the confusion of the "potentially" or capacity with the "Actuality" or working out of the capacity in practice. Leroy (1) Stockton points out the variety of meanings that can be attached to the term 'adaptability' in its general employment with reference to the inorganic and organic worlds, and suggests as an all inclusive definition from this point of view the following:-

"An organism is intelligent when it possess the ability to influence its destiny through the utilisation of an inner, active, non-predictable, selective factor which chooses on the basis of similarity". It is doubtful if this elaboration of Stern's definition is necessary when due note of the word "consciously" in it is taken. (2) Terman's objection to the definition is similar - that it affords no clue for judging the value/

---

(1) Stockton:

(2) Terman: Symposium on Intelligence. Journal of Education Psychology. March 1921.

value of different kinds of adaptation. He proposes instead a definition more restricted in nature when he states "An individual is intelligent in proportion as he is able to carry on abstract thinking". A definition of this nature does not readily meet with general acceptance on account of its limited reference to the conceptual levels of mentality. It is supported in Terman's belief, by the fact that the tests which survive are precisely those which measure abstract thinking, and, as further evidence of the truth he observes, that the races which survive are characterised by the power to think abstractly.

(1)

Thorndike, whose early experimental work found no trace of a general factor of intelligence, persists in denying the existance of a common factor. He states, "This doctrine requires not only that all branches of intellectual activity be positively correlated, which is substantially <sup>true</sup> one, but also that they be bound to each other in all cases by one common factor, which is false. The latter would require that no two intellectual abilities or branches of intellectual activity should be more closely related to each other than to the fundamental function by which alone they are supposed to be related ..... But, unless one arbitrarily/

---

(1) and Thorndike: Educational Psychology. Vol. III  
P.P. 364-366.



arbitrarily limits the meaning of "all branches of intellectual activity", so as to exclude a majority of those so far tested, one finds traits closely related to each other but with their common element only loosely related to the common element of some other pair ....The mind must be regarded not as a functional unit, not even as a collection of a few general faculties which work irrespective of particular material, but rather as a multitude of functions each of which involves content as well as form, and so is related closely to only a few of its fellows, to the others with greater and greater degrees of remoteness". Thorndike's original adherence to the many-sided point of view has become restricted, in the main, to a recognition of three distinct types of intelligence. These are Abstract Intelligence, or the power to handle words, Motor Intelligence and Social Intelligence. That the last named at least, contains factors other than those which are ordinarily designated 'intelligent'. Thorndike would probably not deny, since, in the afore mentioned Symposium on Intelligence, he suggests the unsuitability of separating out intelligence from emotional and creative factors. Regarding the measurement of these factors he says "In general tests which require efficiency in analysing a situation into elements selecting and weighing elements to fit a problem.

But a straightforward information tests is also valuable."

(1)

Meunaun/



Meumaun also criticises Spearman on the ground that he merely translates an arithmetical result into a psychological assumption. He prefers to regard intelligence from a practical point of view. A man's intelligence bears on his ability to create values in life, desire new theories and form new aims. The intelligent individual overcomes difficulties or avoids mistakes in action. Three aspects are revealed in Meumaun's idea of intelligence, a qualitative, the existence of certain powers of thought, an intensive, a certain intensity in life mentally considered, and an emotional and volitional making possible the development of an individual's natural endowment.

The point of view is interesting, because it reveals a tendency which is all too manifest in definitions of mental ability, to widen the meaning of the concept. In the 1921 Symposium on Intelligence both Thorndike and Freeman exhibit a similar general point of view. The former protests against the separation of intelligent from emotional and volitional factors, the latter desires that the meaning be extended to cover "temperament and moral characters as well, such as 'moral control', adoption of mental purpose etc". Therein lies one of the difficulties that have confronted those who are concerned with the understanding of what is termed intelligence, that psychologists/

86

psychologists themselves are not agreed as to the scope of what they are testing and that delimitations of the sphere of mental testings are vaguely conceived. There seems to be no question that definitions such as employed by Maunaun and others, which include the emotional and volitional aspects of the total mentality, are too wide to be acceptable in reference to the tests employed. If the worth of the tests is established this cannot be challenged. What the tests actually aspire to measure is not the total mentality of the individual, inclusive of emotional and volitional factors, but only ability. It is what the individual is capable of doing generally (or maximally) that it is estimated, not what he may do under the force of special circumstances. As Dr. Drever succinctly observes (1) "The acid test of life, leads maybe, to a different verdict ..... As a general rule, mental tests are the tests of capacity and do not profess to be more. Indirectly and incidentally they may throw light on tendency, but they are not intended to do so". So excellent an exposition of the claims of the mental test and of "intelligence" generally, is offered by Kellogg (2) in his chapter on "Intelligence Tests" that/

- 
- (1) Drever: Report on "Psychological Tests of Educable Capacity".  
(2) Kellogg: Mind and Heredity. (P. 123). 1923.

that a ~~qu~~otation, pardonably long, may serve best to leave the matter clear. "What is it that intelligence tests test and measure? We all know that any person's mental make-up consists partly of something he has inherited from, or better, through his parents, and partly of some things he has acquired from his parents and others acting as teachers and perceptions and examples to imitate - or to avoid imitating - as well as from books, observation and experience and from the personal exercise or lack of exercise of his inherited mental faculties. Among those things he has inherited are general mental capacity and certain specific mental traits, which we can group together under the name of intelligence. And there are also emotion and temperament, natural courage or cowardice, aggressiveness or retiringness, born independence of born dependency, born leadership or born following. Now of all those things inherited and acquired, what are those which intelligence tests claim to, and really do, test? Just and only those, but those highly important ones indicated by the name intelligence, those inherited qualities of general mental capacity and specific mental traits which compose what we call intelligence, meaning native capacity for learning by observation, experience and being taught, mental alertness and suppleness/

# ON THE UPPER LIMIT OF THE GROWTH OF INTELLIGENCE.

---

## The Controversy.

---

The controversy regarding the upper limit of the growth of intelligence is of comparatively recent origin. It dates, in fact, from the discovery which arose from the result of the application of tests to over two million recruits entering the American Army at the time of the Great War. The discovery was made that the average mental age of the population, assuming the large male element tested to be representative, was only approximately  $13\frac{1}{2}$  years. Previously it had been more or less tacitly assumed that the average upper limit at which intelligent growth virtually ceased was at 16 years. The Binet-Simon Scale was then young and up to the opening of the controversy a comparatively short time ~~has~~ elapsed for the scale to make itself known to the different peoples. The scale itself was modelled in its best form only in 1911, but since 1905 when it first appeared, psychologists ~~for many~~ ~~centuries~~ interested themselves in its principles, applied and attempted modifications and adaptations of/



of it to suit the conditions of their own environment. Nevertheless, in no country had any degree of finality marked the results and in many countries the adaptation effected was purely tentative. The first radical adaptation of the scale to suit English children, was not, in fact, available till after the controversy had arisen, when <sup>Burt</sup> BINET published his *Burt?* Memoranda in 1921.

So far, then, the main interest of investigators was with the trying out of the scale itself and with the understanding or questioning of the very principles of the scale's construction. Up to that time TERMAN was responsible for the most thorough revision of the Scale. The STANFORD Revision of the BINET scale was probably more thoroughly standardised than any of its predecessors and was extended beyond the upper age limit which the BINET Scale itself had reached by the addition of many new tests. From his test results TERMAN decided to employ 16 as the basis of calculation for the mental ages of all subjects at and beyond life age sixteen. This was equivalent to placing the average limit of intelligence at that age, a belief which he otherwise held.

As against <sup>this</sup> the view the Army results presented/



presented an average limit of mental growth at  $13\frac{1}{2}$  years. From this initial point the conflict starts. On the one hand there is the view of Terman based on his own experience with the BINET-SIMON scale, and, in particular, <sup>on</sup> ~~as~~ his testing of certain adolescents and adults, and to ~~the~~ view he is nearly as staunch as to his belief in the constancy of the I.Q. On the other hand is the view resulting from the American Army tests, which finds a supporter in Doll. At the same time there are a number of opinions, they are little more, oscillating with considerable freedom between the two <sup>extremes</sup> ~~es~~hems, now edging towards one, now towards the other, and careful always to go with the balance of general opinion.

#### GREEN.

---

Entirely opposed to either of these views and to the sundry qualifications of them is the view of the late Professor GREEN <sup>(1)</sup> of Manchester, insisting that it is the height of absurdity to claim that intelligence ceases to grow on the average either at  $13\frac{1}{2}$  or at 16 years. Incidentally, this is also the view, though based on less scientific reasoning, of the man in the street. It comes as a mental shock to individuals that their intelligence should cease to/

*Sheffield*

---

(1). GREEN: Journal of Experimental Psychology 1922.

to grow after sixteen. They tell you with emphasis that you talk nonsense and that they are far more efficient now than they were at sixteen or so.

BALLARD has replied to such opinion in an attempt to answer GREEN'S objections, by saying that all is true, but it has nothing whatever to do with the subject of intelligence pure and simple. As a man gets older, he feels more at home in the world, and accommodates himself with less effort to his surroundings. But he does not on that account expand or develop further his intelligence.

Generally speaking, it is safe to suggest that there has existed a too optimistic view of the innate ability of man, based perhaps on the brilliant performance of the mentally alert and due also to the fact that the lower intelligences were permitted to sink or swim without the community's interest being attracted to the problems of their existence beyond the mere attempt to obliterate vice and crime by artificial measures. But a general review of life is itself enough to convince the eye witness that the theory of the psychologist is not devoid of reason and that the undue failure to recognise individual differences, and in particular, individual/

individual possibilities or limitations have been at the root of much trouble. If there are square pegs in round holes, there are also little pegs in big holes and vice versa and maladjustments due to unsuitable accommodation.

But BALLARD'S reply does not meet the whole of GREEN'S position. GREEN claims that the investigations of intelligence are based on two assumptions :-

- (1) That there is something in the nature of mind energy which enters in varying degree into all controlled behaviour from the simple repeated movements, through sensory discrimination to the level of analogies, and synonyms.
- (2) That a hierarchical system of correlation, such as found by BURT in his early investigations at Liverpool, is evidence of the existence of a central factor.

Now strictly speaking, there is no such uniformity in the ideas of psychologists regarding the nature of intelligence, as has been seen, but GREEN next goes on to suggest an alternative starting point. It is better to begin with the biological point of view and "regard intelligence as an integrative function of mind, which in the last resort finds expression in adaptive behaviour". In this way intelligence is a simple function "operating at/

at call but with varying efficiency". It ought to unify experience, but it operates at different mental levels. Mind never works in a vacuum. It may create its own world.

Two things follow from this view.

- (1) That intelligence is not a summation of parts or a multi-dimensional entity capable of resolution in certain directions.
- (2) Intelligence is correlative with the universe in which it works. As the universe develops, intelligence develops. They condition one another and intelligence is revealed by its universe of operation. In other words, we cannot measure a potential intelligence.

On account of his theory GREEN cannot believe that intelligence ceases to grow after 16 years and even if it happens that a boy of 16 does as well in tests <sup>as an adult.</sup> (as e.g. BALLARD'S Absurdities) GREEN believes that his position is still correct, since he would not admit that the detection of absurdities was a test of intelligence. The reason why the boy of sixteen succeeds as well as, if not better than a man of forty, is that the test is suited to the environment of the former and alien to the environment of the latter. If the boy were put in the adult's sphere he would succeed as badly.

The position of GREEN, so outlined, contains/

contains certain truths, but as the factors<sup>0</sup>~~are~~ are interpreted by GREEN, these are not altogether clear. The instant suggestion that arises in the reader's mind is that GREEN is not primarily a psychologist. He was, in fact, primarily an educationist. His criticism of the psychological views is based on a failure to inquire into what the psychologist is really testing. The biological outlook is probably sound but it has for the psychologist the demerit of not being psychological. However much the psychologist finds community with the ideas of the biologist it is not on that account permissible to substitute the biological for the psychological point of view. Confusion in interpretation is the natural outcome of such a proceeding. A parallel case is revealed with regard to STERN'S famous definition of intelligence. It is stated that intelligence is more than adaptation and so there comes about a gradual dethroning of the long spoken of favourite. But STERN did not deny that intelligence was more than adaptation. He said that intelligence was 'conscious' adaptation, an entirely different thing, and incidentally, with the merit of being psychological and having a psychological reference. So it/



it is with GREEN'S position. His fundamental analysis leads him to neglect the fact that the intelligence tester makes no claim to assess the total make up of an individual. Far from doing so, he rather bewails his incompetency, as yet to measure one aspect of the total mentality.

There is no denying the truth of GREEN'S assertion that as the environment develops so does the intelligence and as the intelligence develops so does the environment, that they are in fact correlative biologically. But what GREEN means by intelligence has a much wider connotation than what the intelligence tester means by intelligence. Not once but many times has it been expressed that the word intelligence is unsuitably employed to describe what the mental tests test, because of its relationship with the term "intellect" and with the same term employed in the study of animal psychology. BURT and SPEARMAN are most careful to avoid confusion the one by employing the term <sup>"</sup>innate all round mental ability<sup>"</sup>, the other by refusing to identify his "g" factor definitely with general intelligence as ordinarily understood. In reality, it is possible to acquiesce in GREEN'S outline with compromise.

To/

To reduce his position to terms of simplicity that are almost childish it is this. Intelligence and environment exist and they are found together, or are correlative. Whenever there is intelligence there must also be environment. In the first instance their relationship is a very simple one since that is all that is necessary. As GREEN would say "A simple intelligence is correlative with a simple environment". But, gradually, both become more complex and in the process they are never one before the other or one after the other. So the development goes on, until as in the modern environment great complexity characterises both. What is significant is that the process is so slow as to require more time than can be thought of even with difficulty, but the selection of the fittest and the elimination of the lowest, if these are accepted as probable processes, are virtually working themselves out, though their much protracted influence evades observation at any point in the series. Even Professor GREEN, however, might have been willing to admit that at any cross section of the period of time it could be found that variety reigned supreme within the species. There would exist intelligences that/

that were complex enough to survive ultimately, and intelligences that were bound to be wiped out, environments of a more simple nature and environments of a more complex nature correlative with the former. Further, at any point of time, such as a day, or a year or a decade, which is small biologically speaking, it is not inconceivable that the blending of environments and intelligences might not be so simple as GREEN would like to think when he matches simple with simple and complex with complex. If this suggestion seems fantastical from the biological point of view it is because of the difficulty of taking and observing a cross-sectional fragment. It is nevertheless borne out by the facts of everyday life at any point of time which the psychologist reviews. The real objection to this outlook, however, lies in the width of meaning which covers the terms intelligence and environment and the width of relation between them which is recognised. The term intelligence applies to the whole adaptative mechanism of the individual. But as the mental tester uses the term, in the restricted sense of "innate mental ability" there is no such comprehensiveness implied. Intelligence is not always correlative with/

with environment because circumstances are such that innate ability is not always given the best scope for its development. Only the brilliant transcend the handicaps of a poor "environment"; even the average tend to sink in it. To use BALLARD'S phraseology the 'potential' mental ability can differ from the 'actual' mental ability seen in operation. Intelligence in the restricted sense of "innate mental ability" is not the only factor, moreover, in the adjustment of an individual to an "environment". Emotional and volitional factors contribute their share and just how great that share can be is probably not generally recognised. It all reverts to the obstinate and persistent belief of the populace that the mental tester claims to have labelled the individual for all time and in every respect when he has given him a short intelligence test.

Such then are the weaknesses of the argument which GREEN propounds, firstly, that intelligence as he employs the term, does not apply at all, since it is not what the intelligence tester means but has a biological significance and secondly that intelligence in the narrow sense is not the sole factor in the mentality of the individual and intelligence tests claim no more than to isolate one factor and to test it by the best available means.

What/

What the mental tester means by intelligence is cut and dried in the definition, however inadequate, that he employs. The claim attached is that the so called "mental ability" can be measured apart from the effects of school training or other educational influence. Consequently the argument of GREEN that the man of forty will fail in the environment of the child of sixteen ought not to hold. As well might it be applied with regard to the environment of the four year old as contrasted with that of the sixteen year old. It is however, not strictly possible to be dogmatic on the point, since among those who have laboured hardest with tests is the greatest doubt apparent. Dr. BURT'S Memorandum finishes up with an assertion that BINET test scores are not uninfluenced by external influences and that a child of ability but devoid of opportunity would fare badly on that account. The statement of YERKS in regard to results from the application of the Point Scale to older subjects again constitute a warning. Further still, it must be confessed that there is some truth in GREEN'S suggestion that not all tests employed are adequate measures of intelligence. But these matters awaiting decision, do not effect the general argument against GREEN'S position./



position. The controversy regarding the upper limit of the growth of intelligence is not substantially influenced by GREEN's opposition.

# IN SUPPORT of the SIXTEEN YEAR LEVEL.

## TERMAN.

In constructing the Stanford Revision of the BINET SCALE, TERMAN standardised his higher age tests chiefly on the basis of results from about four hundred adults. These were some 32 High School Students, 60 business men, 150 migrating unemployed chiefly between 25 and 40 years, and 150 juvenile delinquents mostly between 14 and 19 years. The tests in the scale extend to the 'Superior adult' level (18 years or so). From the result of these adult subjects he concluded that the average mental age for normal adult subjects was 16 years. It is questionable whether TERMAN'S conclusions here are justified, but, as he himself has more recently agreed that the figure is somewhat too high and that 15 years would more nearly represent the upper limit for average intelligence, there is no need to be over-fastidious on this account. Regarding the growth of intelligence <sup>(1)</sup>TERMAN writes "We are warranted in believing that general intelligence practically ceases to develop by the age of 18 or 20. The/

---

(1) TERMAN: Journal of Educational Psychology. Sept. 1921.

The mental age of high grade morons appears to change little after the age of 14 or 15 years. With normal children development continues a little longer, though at a decreasing rate. It is practically certain, however, that growth of intelligence comes to a standstill somewhere in the later years of adolescence and that the cessation is gradual rather than sudden". Further Terman recognises that his view differs from those of others, such as the view that intelligence develops by "leaps and bounds". In addition, he suggests that, as far as results go the "much talked of adolescent ~~spurt~~ <sup>spurt</sup>" looks like a myth. It is hardly necessary to add that when Terman speaks of the limit of growth being about eighteen or twenty years he is not referring to the aver. upper limit but to the maximum. It is owing to a failure to observe the distinctions between the absolute final limit of growth and the "average" upper limit of growth that much confusion has arisen in the criticism of Terman's view and the views of other writers.

Regarding the rate of growth Terman maintains that it progresses at a constant rate for all degrees of brightness and dullness (with the exception of imbeciles and idiots) but, as seen above that/

that mentally defective children fall off at earlier ages than do normals, who on the average progress beyond the defective limits. DOLL thus, misinterprets Terman when he affirms that Terman makes no reservation regarding the age of arrest of mentally defective children as also when he suggests that Terman proposes the same rate of growth for all degrees of ability.

In general criticism of Terman's position DOLL, the supporter of the  $13\frac{1}{2}$  upper limit level, refers to the 62 "normal" subject which Terman employed from schools and business spheres, to obtain the norms. above 15 years. His complaint is that they were supernormal being high school children and successful business men. In the Grammar School Groups which Terman tested the median mental age was only 14 years, but Terman declares that they were subaverage, being retarded for their age. DOLL suggests that a combination of these results with the former would yield a satisfactory average of about 14 years. Then, commenting on Terman's later work<sup>(1)</sup> DOLL suggests that by combining/

---

(1) Terman: The Intelligence of School Children.

combining the averages which TERMAN obtained from College Students with the averages obtained from many other types, the mental age of 13 would be obtained and would be nearer the truth. It is characteristic of DOLL as will appear later, that he does not give only one conclusion.

TERMAN'S position was more or less accepted till the method of group testing with a time limit to performance was employed in sorting out the recruits of the American Army, to fill the posts to which they were best suited. The Army Test, based on work which Otis was engaged on, was the most extensive investigation that has been conducted, nor has it been equalled since. In this investigation men of every social status and great range of ability were given the tests so that it seemed as if here, if anywhere a claim, that the results were representative of the average population, might be supported. But the suggestion was not mutely accepted by those believing in the 16 year level and if the suggestion of a cessation of growth at 16 had upset the population this latest result must have astounded them.

TERMAN<sup>(1)</sup> against whom the new findings reflected/

---

(1) TERMAN Journal of Educational Psychology.  
Sept. 1921.



reflected more, refused to accept the results. He stated certain reasons for the non acceptance of the view that  $13\frac{1}{2}$  represented the average upper limit of mental growth. These reasons were:-

- (1) That his own results go to prove a higher level, and, even if the validity of these is in question, the Stanford Revision of the BINET SCALE is probably somewhat too difficult at the upper ages.
- (2) The men were not representative of the draft army as a whole. A disproportionate number came from the South States where intelligence is lower than in the Northern States and Western States.
- (3) The draft army itself was not representative of the male population between 21 and 31 years. "The cream of the Country was excluded".
- (4) The conditions of testing were far from ideal.
- (5) An abbreviation of the STANFORD BINET SCALE was used, and while the abbreviation yields scores which correlated very highly with the scores of the entire scale, it can be shown that with adult subjects they tend on the average to run slightly lower.
- (6) Untrained examiners were employed.

Writing more recently, KOHSS<sup>(1)</sup> like TERMAN a supporter of the 16 year level gives reasons for rejecting the other view. Such of the reasons as differ from those of TERMAN may be added at the point. They/

---

(1) KOHSS. Intelligence Measurements.

They are summarised below.

- (1) That the Army test, together with the conditions under which it was taken "in and of itself requires an adaptation to a new situation, which is quite complex". This is an undesirable factor since all subjects will not adapt themselves equally and there will be a drag, especially on subnormal intelligences.
- (2) While it is generally assumed that ability in a group test will correspond to ability in a BINET individual test, close analysis will reveal that "Whereas a high score in a reliable group intelligence test is unquestionably indicative of that mental level, because superior scores cannot be obtained by accident or haphazard the same does not hold for his scores". Any number of factors may have entered in to affect the lower scores and the chances are the lower the score is the greater will be the number of factors. He adds a figure to show that this "phenomenon of progressive retardation would yield to the median performer who obtained an intelligence rating of 16 on the BINET, an intelligence rating of barely 14 in a group test and thus in spite of the fact that the result would yet yield a correlation of unity.
- (3) The Army Alpha over-rates reading ability.
- (4) There is over emphasis on the Speed element in Army Alpha.
- (5) Whereas it is generally believed that approximately 2% of the whole population is mentally defective, the army data would, employing the same criteria, rate 30.3% mentally defective. Employing eight years instead of 12, the army results would give 2% mentally defective, but the question now arises of individuals of mental age eight and mentally fit to manage their own affairs/

7000

affairs. The problem here reflects on the proportion of feeble-minded who are in institutions and it also reflects on the employment of the existing I.Q. method of classification of intelligence.

- (6) There was a constant need for supplementing the BINET tests in examination by group tests. These required additional directions which probably were not equal to all tested.
- (7) For a check on the testing the STANFORD-BINET age of 653 draftees from 9 various camps were employed. The assumption that these were representative of a random sampling is only questionable. (KOLIS enters into details, which show that the scores were lower)
- (8) The Pragmatic Test has supported the 16 year level.
- (9) Through the employment of group tests the possibility of relationship between the individual and the tests was cut out.
- (10) Numerous factors of personality and temperament undoubtedly had an effect on the results.
- (11) So far, statistical sanction has indicated no error in the ~~ts~~sixteen year old criterion.
- (12) Even the BINET tests probably are not adequate for determining adult mental age.
- (13) On the basis of any statistics only 15% of the draftees examined ~~measured~~ over the years of age. ~~Were~~ 85% then below average?
- (14) If the exemption of the feeble-minded counter-balances the elements of the superior men, then the results might be valid. This has been claimed to be so, but, in fact, only the lowest were eliminated.
- (15) A satisfactory standard of the average mental age of adults should include statistics gathered from women.
- (16) Common/

(16) Common sense objects to the 13 year level.

(17) Lastly, YERKES himself points out the danger of taking the figure too literally. They involve an error dependent on the fact that the men were tested by Alpha and Beta and not by a mental age.

So much for KOHS' evidence. It is long and it seems that its clarity has at times been sacrificed to its exhaustiveness and that KOHS could have expressed the essentials in briefer space. Some of his objections are founded on fact and are worthy on that account. Others, such as the comment on the emphasis on reading ability and temperament apply to nearly all test situations and are not to be regarded as exclusive to the Army Alpha testing on that account. The point treating of the relation between tester and tested is of doubtful significance in this matter. It is not undesirable that there should be no relation between the two, especially when they are adults. Where KOHS confirms Terman's beliefs he is interesting but the contribution is chiefly interesting for several suggestions which it made regarding not only the Army Results but the method of approach to the whole question. If the theory of "progressive reduction" holds good, then here is another argument/



argument to add to these already existing against the employment of group tests for solving the problems of mental growth. Or again, if the factor of speed is a real handicap to <sup>some</sup> ~~solving~~ subjects then speed tests must be regarded as inadequate. This remains to be proved but it is probably true that, in conjunction with the far from perfect conditions under which the testing was undertaken, the factor of speed constituted a real difficulty and a real handicap for certain types of intelligence.

A further contribution to the whole argument, presented by LINCOLN (1) does not at first sight bear the impress of a favourable support of Terman's view. LINCOLN starts by questioning all Terman's arguments with the single exception of the first, relating to the STANFORD REVISION. He doubts the validity of Terman's statement that the draft was not representative of the army as a whole. "Evidence against Terman's view, though slight, is cumulative". The assertion that the draft was not representative of the whole population, but that the cream of the country was excluded, may also be questioned. The statement can neither be proved nor disproved, but among those excluded were "dependent", mostly married men. LINCOLN suggests that early marriage is not a sign/

---

(1) LINCOLN: Journal of Educational Research.  
Sept. 1922.



sign that individuals are mentally superior. The most intelligent men are for the most part, engrossed in other matters between twenty and thirty years. Aliens, men excluded for industrial and occupational reasons, men physically and morally unfit could none of them be regarded as the cream of the Country. Further, though College Students and graduates, being volunteers, were exempted, these were not disproportionate in number. LINCOLN suggests that in place of multiplying reasons of this nature, TERMAN should rather have turned to the schools themselves and noticed that children surpassed the army average scores in the Alpha Examination. Referring to DOLL'S investigations at Franklin he points out that the upper averages obtained are low because the best children had probably gone to the High School. He reports an investigation carried out with High School children in Madison, Rockford and Sioux City carried out by SYLVESTER and DOBBS. (School and Society. Oct. 1919) The results are in terms of the Army Alpha scores and are reported below.

Table/

	M A D I S O N.		R O C K F O R D.		S I O U X C I T Y.	
	MEDIAN SCORE	MENTAL AGE	MEDIAN SCORE	MENTAL AGE	MEDIAN SCORE	MENTAL AGE
Senior	121.2	17.4	123.3	17.4	137.6	
Junior	122.4	16.6	119.7	16.4	130.2	16.6
Sophomore	109.7	15.3	107.0	15.3	121.1	16.0
Freshmen	96	14.3	94	14.5	107.9	14.8

At first sight it might appear strange that LINCOLN is on very dangerous ground in using High School children for the problem. Their average mental scores indicate their superior mentality at all stages. But LINCOLN pertinently asks "Can selection alone account for the 14 year old scores being 50% above the Army Alpha general average and for the 16 year olds being 100% above the Army average? (Roughly, 60 points is the Army Average)

Again, a report of another investigation in which High School pupils are tested by KOHS and PROCTOR shows that these investigators find a lower mental age to approximate to the Army Alpha Scores than the Army Alpha suggests. The table they give is as follows:-

Alpha Scores	KOHS	Army Alpha
45	12	13.0
50	12.4	13.3
55	12.8	13.6
60	13.0	14.0
65	13.4	14.3
70	13.8	14.6

The/

The exact age at which these two investigations would indicate a cessation of development is not indicated. They are content to disprove the army Alpha results from practical testing experiences. The inadequacy of these experiments is apparent but they indicate a need for further elucidation of the matter. Probably the experimenters intend to support TERMAN'S view of a sixteen year old level since they reject the Army findings.

In addition YERKES Point Scale results point to development up to about 15 years with little or no improvement thereafter, while KOHS' mental age equivalents of total scores from the Block Design Tests indicate a levelling at sixteen years.

BURT'S Group Test (1) norms reveal a slowing down after 13 years. Regarding the placing of the upper limit of intelligence growth at the place indicated by the Army results, he has this (2) to say "With this extreme deduction my own results hardly conform. Apart from the immense accession of acquired knowledge and skill, setting aside too the/

---

(1) BURT: Mental and Scholastic Tests. p. 222

(2) IBID p. 244.

the gradual emergence of new powers of character and feeling which the process of adolescence seems to confer, there can indeed be little doubt that, after the age of leaving school, the further development of natural intelligence is in most persons far smaller than is commonly thought. Nevertheless, the opposite assumption that intelligence grows by equal annual increments up to the beginning of puberty and then abruptly ceases is equally mistaken..... Meanwhile, those who consider that ..... my upper limit (16) is too high..... might be willing to accept as a temporary compromise an age of 15 for the virtual cessation of normal mental growth<sup>7</sup>

#### AGAINST Terman's View.

DOLL has already been indicated to support the results of the American Army Tests and the placing of the average upper limit of mental growth at  $13\frac{1}{2}$  years. His view is based on experimentation with tests of the BINET order as applied to subnormal children. In addition to actual results he offers his personal "Opinions" on the matter. The interest of the study lies in the point of attack. The whole work is left over for minute analysis later.

of/



Of the smaller investigations that of DEARBORN (1) merits some discussion. DEARBORN states a preference for the lower level, but his actual conclusions scarcely bear out his support. He presents data from the testing of unselected groups of 14 and 15 year old children and "tests the assumption of a lower mental age for the average adult (than Terman's) by applying it to data secured from the testing of practically all of the children from three communities, between the ages of 8 and 13 years and those who remained in school after that age".

In the problem of fixing the average upper limit for the growth of intelligence he recognises two difficulties. The first is due to the variability of individuals. If we recognise adults of mental age two, we must recognise the possibility of variability as great towards the other end. Hence our average does not preclude the possibility of some progressing mentally beyond the twenties and even farther. But as the adult idiots do not represent more than a very small percentage of the adult

---

(1) Journal of Educational Research.

adult population, neither will the exceedingly superior. (This ought to act as balm to those unwilling to admit that their mental growth ceased in the adolescent period!) The final status of the individual "variant" depends on three factors viz: initial ability, rate of growth and duration of development. Perhaps even a fourth might be added—the factor of environment. There is evidence, in DEARBORN'S opinion, of correlation between them, so that an individual of special ability develops faster and during a longer period of time than an individual of average ability.

The second difficulty arises when it is seen that mental age increments seem to increase up to the age <sup>of  $13\frac{1}{2}$</sup> , and this makes it difficult to accept  $13\frac{1}{2}$  as the level, since presumably growth does not cease abruptly. DEARBORN quotes the norms (below) of his own group tests to prove this gradual decrease of growth.

AGE.	$8\frac{1}{2}$	$9\frac{1}{2}$	$10\frac{1}{2}$	$11\frac{1}{2}$	$12\frac{1}{2}$	$13\frac{1}{2}$	$14\frac{1}{2}$
SCORE	25	37	49	51	73	85	97
NUMBER of CASES	400	429	395	413	379	412	337

The/

The  $14\frac{1}{2}$  group are selected so that the increase, which is equal to all previous increase, is really a sign of decrease in average growth since it is partly due to growth and partly due to selection. If DEARBORN reasons from this, as he seems to, that average mental growth ceases at  $14\frac{1}{2}$  then it might be asked where his prior recognition of the gradual decrease of average growth comes in. From the table  $14\frac{1}{2}$  seems to represent the point at which decrease first sets in on the average and it is not unreasonable to assume that it becomes still lower with another year and probably improvement then becomes negligible. Incidentally it may be noted that DEARBORN is keenly alive to the inadequacy of group tests method which conceals individual differences in ability and in rate of growth. He recognises also the possibility of variation in the form of growth curve due to differences in the nature and difficulty of the tests employed.

The other part of the investigation involved the testing of children of 14 and 15 years, attending Regular Schools and Continuation Schools in Massachusetts. Attendance at one or other type of school/

school is compulsory between fourteen and sixteen years. Continuation School children represent unselected subject material which usually drops out of the school life at fourteen years. It is the study of these children which would be most advantageous for the fixing of the average upper limit of intelligence growth.

The results of DEARBORN'S investigations are placed below. DEARBORN adds distributions to show the great similarity of form in the distributions of Regular School and Continuation School results. He explains the increase in the average scores of the 15 year-olds over the scores of the 14 year-olds as being in some part due to elimination of pupils to Continuation Schools.

Table/

MEAN SCORES OF MENTAL AGES OF PUPILS. FOURTEEN AND  
FIFTEEN YEARS OLD.

	In REGULAR SCHOOLS		In CONTINUATION SCHOOLS		In ALL SCHOOLS	
	Age 14.	Age 15.	Age 14.	Age 15.	Age 14.	Age 15.
Average Scores	119	129	81	89	110	111
Mental Age	14.5	15.1	11.10	12.4	13.9	13.10
No. of Pupils	342	249	112	208	454	457



DEARBORN does not accept his own results (seen in the column headed "All Schools"), that intelligence on the average practically ceases to develop at fourteen years, since many of the children tested were of foreign parentage and tested lower than the ordinary children would have done. He suggests a slight heightening of the figure to  $14\frac{1}{2}$ , and offers a further justification for the suggestion. By using the figure fourteen as the denominator in the calculation of the I.Q.'s of the subjects tested, from the Regular Schools, he obtained a gain in the I.Q.'s for the fifteen year olds, a finding which he believed to be necessary since the fifteen year olds remaining in the Regular School were almost certain to be superior in average ability. If on the other hand he employed the figure sixteen as the basis on which to calculate the I.Q.'s at both ages, the fifteen year olds obtained lower I.Q.'s than the fourteen year olds.

The study is unfortunately limited in scope but it seems in reality to have a leaning towards the fifteen year old level suggested by TERMAN and others rather than to support the view resulting from the Army testing. A further extension of the investigation beyond fifteen years might have shown that average/

average growth was still preceptible. The combination of the results from both schools conceals much, especially the considerable difference in the average scores for the Continuation Schools fourteen and fifteen year olds.

Within the last few months another group study has appeared (1) which is discussed here on account of the affinity it bears to the study to be presented later, and to reveal certain inadequacies in interpretation of the results.

R.R. DOBSON'S investigation had a three-fold aim.

- (1) To determine if written tests could select the brightest children in the school.
- (2) To determine by means of such tests whether there is a limit to the growth of intelligence and at what age it is reached.
- (3) To determine which of the particular tests employed depend on innate ability and which on school environment and teaching.

He employed in the investigation Grammar School children, both boys and girls. The boys' age ranged from eleven to 13 years, the girls from 16 to 18 years. Men in attendance at an Elementary Teachers' /

---

(1) DOBSON: British Journal of Psychology. Oct. 1924

Teachers' Training College, University Students, Honours Students reading for degrees. University Lectures and Professors, were also tested. The tests were prepared by Burt on Civil Service lines and were contained in book form. They comprised synonyms and opposite, analogies, mixed Sentences completion and reasoning tests. They were time limit tests which required only under-lining for solutions. The scores obtained were correlated with other estimates of intelligence. In the table below the essential results are quoted. The total score available was 189.

BOYS RESULTS.	
AGE	SCORE
11 - 12 =	84.2
12 - 13 =	91.1
13 - 14 =	100
14 - 15 =	109.5
15 - 16 =	112.3
16 - 17 =	125.2
17 - 18 =	135.4
TRAINING COLLEGE MEN	
Av. Age $19\frac{1}{2}$ =	116.3
HONOURS WOMEN =	
	130.6
and so on.	

DOBSON'S conclusions from the full table of results, which is only partly reproduced here, are that Grammar School boys improve up to 17 or 18 years and that the University Students make even higher score than the others. The girls test lower than the boys. DOBSON expresses his conclusions so "There is an increase in age performance in all tests up to the age of seventeen years. There is little or no improvement after that age. Judged by the total score it appears that, for the great number at any rate, intelligence apart from experience ceases to grow except among people of exceptionally high ability". Further from the scores obtained by the "two year Students" he suggests that the best boys and girls from the secondary school do not go in for elementary teaching.

The conclusions of DOBSON cannot be accepted at their face value. There is no evidence that the subjects tested from age to age were average subjects and not selected and there is positive suggestion that they were. They were secondary school pupils it appears and it is surprising how small is the total percentage of all children in the secondary schools. Moreover, they are not only few but they are/

are generally able<sup>1</sup>, at least beyond the age of compulsory attendance. The results are singularly parallel to the results obtained in the following experiment, but the interpretation differs. The results themselves offer the clue to detection of selection. The average scores of the boys indicate a fairly steady increase from  $11\frac{1}{2}$  to  $12\frac{1}{2}$  and from  $12\frac{1}{2}$  to  $13\frac{1}{2}$  and during the next interval. But it is significant that after 15, at the point where selection would be expected to operate with children of this type the average increase is not half as great as it was before. Between the 15-16 and 16-17 groups the average growth is recorded as 12.9 units, which is greater than the average growth of nine units in the two intervals between the 13 and 15 years, and therefore obviously due to the operation of selection. In spite of this, DOBSON finds the average upper limit for the growth of intelligence to be situated at 17 years. Now, the immense increase in average score between the 15-16 and the 16-17 level of age is presumably due to selection. It is preceded by an average increase in score from  $14\frac{1}{2}$  to  $15\frac{1}{2}$  which is very slight, so slight in fact as to raise the question whether selection did not operate before that age but concealed itself in/



in combination with selective factors. Further evidence of a 17 year old level being too high for the average appears when the Training men's results are noted. They compare only with the 15-16 years group at the Grammar School, and they are doubtless selected themselves. The same applies to the results from other of the adult groups tested. Until or unless there is added supplementary information regarding the possibility of selection being at work in the middle of the Grammar School Period, the means of explaining the abnormally large average growth between 16 and 17, lies heavy on the investigator's shoulders.

## THEORETICAL CONSIDERATIONS OF GROWTH.

---

Qualitative differences between individuals and between children and adults were recognised long before an attempt was made to measure them in quantitative terms. Parents would rhapsodise on the brilliance of the child or bemoan the stupidity of that. That a village had a village idiot and other peculiar characters was no unknown occurrence. Moreover, it was generally recognised that the average child of four years was less intelligent than the average child of sixteen years and so on. But in spite of the recognition of individual difference at any stage of growth, and of differences between individuals at different points of the growth period, it was left to Binet to devise the first scale with which to attempt a measurement of the existing differences.

Two assumptions underlie the use of mental tests. The first is that intelligence can be measured apart from knowledge, the second, that the growth of intelligence follows an orderly development from birth to the time of maturity, whenever that may be.

In/

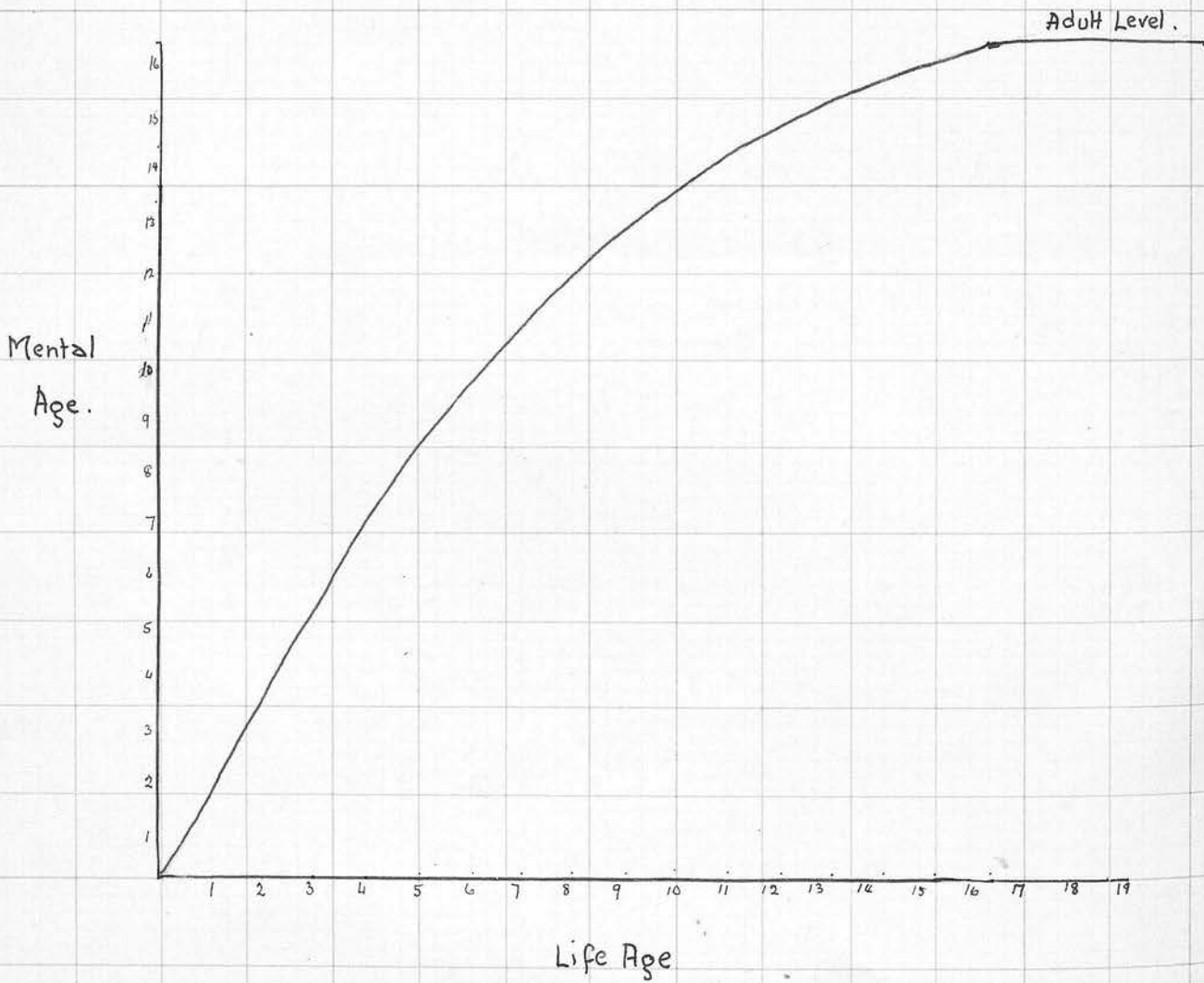
In confirmation of the former belief stands the whole body of evidence regarding the correlation of test scores with various estimates of intelligence as well as the practical outcome of the testing itself, as everywhere evident, even with tests which, as yet, make no claim to perfection, but which are gradually being shaped to better forms.

No claim is made that the scale employed is the best conceivable instruments, that could be discovered for the purpose. A deficiency in test material is especially characteristic of the upper years, where normal school life is finishing. But the outlook as regards mental tests, though it has not yet passed through its period of infancy, is promising of good future developments.

As regards the latter assumption, it is generally conceded, and testing has tended to establish as a fact, that mental growth tends to be a gradual process, not given to leaps and bounds, to periods of stagnation and periods of extraordinary "spurts", and that these characteristics if found, are exceptional and due to particular circumstances.

Since/

# Theoretical Mental Growth Curve



Since mental growth can be measured and follows a regular course of development it is possible to construct a theoretical curve of mental growth in the same way as we would construct a curve of physical growth. The curve is generally believed to be similar in form to that obtained from physical measurements, such as height or brain weight, or any other trait. Such a curve exhibits the form shown opposite. Growth proceeds with regularity during the early years and towards adolescence exhibits a gradually slowing down movement reaching a final "upper limit" somewhere during the adolescent period.

The Binet-Simon Scale, empirically constructed so that the medium mental age corresponds to the medium life age, and based on the assumption that increments in mental age from year to year are equal in amount, yields a curve which is virtually a straight line. Its equation is  $y = x$  up to the limit of the scale Binet's concern was with mentally defectives and the exigencies of his duties neither compelled nor enabled him to carry his scale beyond twelve years satisfactorily.

Goddard's Revision of the Binet Scale,  
based/



based on the principle of locating the different tests, a point which BINET had overlooked shows an efficiency limit at 10 years. That is to say the equation it yields is  $y = x$  up to 10.

Similarly, TERMAN'S scale, constructed on the plan of deliberately securing an arrangement of tests whereby the median mental age of unselected subjects would correspond with the median chronological age, and supplying a deficit in the other two forms of the scale by the introduction of many new and excellent tests at the upper end, shows an efficiency limit at 16 years. That is to say, the equation is  $y = x$  up to 16. TERMAN was handicapped by no such purpose as BINET in constructing the scale but experimentation with the Stanford Revision proves it to be somewhat unsatisfactory at the upper end. This suggests either a difficulty of devising tests at the upper levels, or an early cessation of the development of intelligence. In calculating - I.Q.s, TERMAN found it convenient to adopt this as the basis of calculation of all mental ratios appertaining to "adult" subjects.

DOLL has calculated from the tabulated results of GODDARD and TERMAN the exact form of the curves/

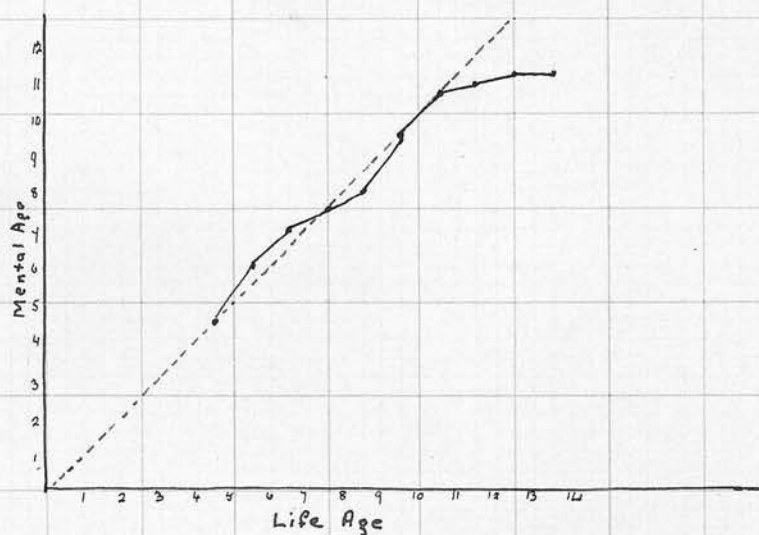
# Binet

(Theoretical)  
Curve



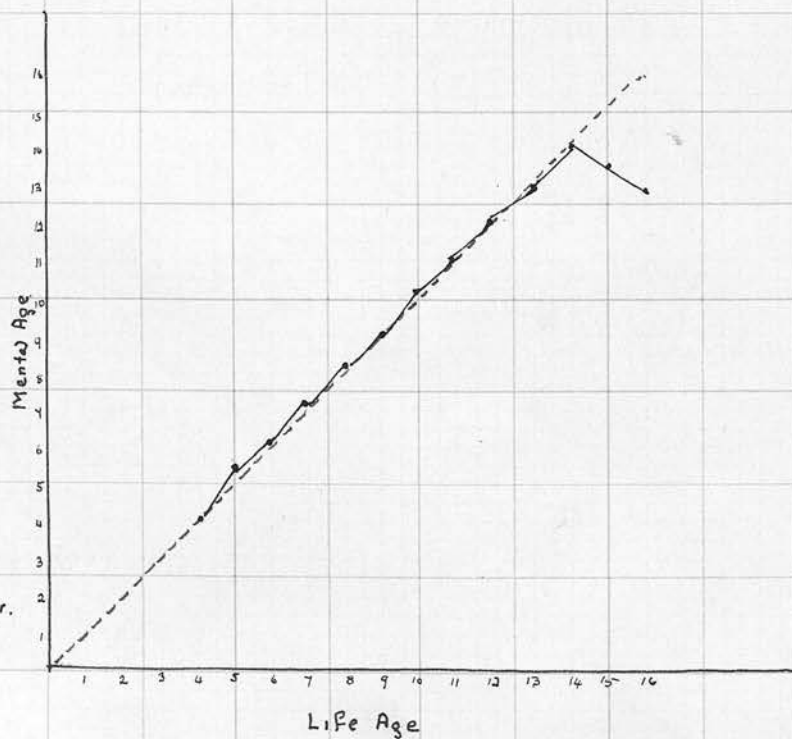
# Goddard

(From testing 1597 school  
children)



# Terman.

(school children & adults)



Scale. 2 units horiz. = 1 year.  
2 units vert. : 1 mental yr.

Extensive use of  
 curves obtained from their scales. A facsimile of this is placed opposite. The curves indicate what experimentation from time to time has revealed, that the GODDARD SCALE is too easy at the lower end and inadequate at the upper end, and that TERMAN'S SCALE is efficient up to 14 years at least. The retrogression thereafter is explained by TERMAN as being due to the testing of sub-average subjects. But the TERMAN higher tests are really too difficult for the ages to which they are assigned.

As before suggested, the true mental growth curve is probably a modification of the theoretical BINET curve, and is probably analogous to the curves for physical growth. But whatever form the true curve takes it reveals always a slowing down towards the adolescent period which gradually becomes more and more evident so as to suggest an absolute average limit in the adolescent period. The result of investigations, hitherto undertaken, to determine this point are as conflicting as opinions on the subject itself.

The age of final mental arrest, as indicated on a curve of mental growth depends on the initial ability of subjects, on their rate of growth and on the duration of their growth. The factors do not permit themselves to be completely separated one from another; they act together to produce the final result. The present stage of development of mental examining is such that a complete knowledge of their relationship is not possessed. Were it possessed the solution to the problem of mental arrest would be forthcoming.

Theoretically not one curve is required but several for the different degrees or "levels" of ability. It is customary and practically it is necessary to separate idiots from imbeciles, imbeciles from morons and so on. A theoretical curve of growth will then be required for each "level" of ability. But it is not to be forgotten that the several curves exist for convenience only and are absolutely artificial constructions. Practically, it is obvious that there is not a clear line of cleavage between imbecility and moronity any more than between moronity and average ability; they shade into one another at some point as border-line cases/

cases show and as qualitative studies make clear. This being so the curves require to be carefully employed with a full recognition of their artificiality.

The matter of constructing theoretical mental growth curves is not by any means resolved when the different "levels" are recognised. The question of rate enters to complicate their formation. At the present time nearly every possible conclusion regarding rate of mental growth is suggested by one investigator or another. Regarding progress in mental development as regular, at least on the average, there exist the possibilities of all "degrees" developing at the same rate, all at different rates, or some at one rate and some at another, it may be proportionately to mental ability. Further, progress may not be spasmodic yet unequal at different stages of growth. If the so-called "adolescent spurt" in physical growth is paralleled in mental growth it is an indication of this irregularity.

One of the most controversial topics in mental testing has a peculiar relationship to the subject of rate of growth. As a means of measuring mental/



mental retardation, BOBERTAG and others substituted for BINET'S 'difference between mental and physical age' method the I.Q. or 'ratio of mental to physical age' method. STERN, TERMAN and others have not only recognised the superiority of this over the old method, but have also emphasised its predictive value. If an I.Q. say of 75 can be regarded as representing a ratio of mental to chronological age at all stages of growth then the limits of an individual's capacity can be predicted. If the constancy of the I.Q. were established this predictive value would be enormous. Unfortunately at the present time there is no general agreement that the I.Q. is constant. It is not proposed to enter here into the details of the controversy at the present time, but only to indicate a matter of vital importance in the construction of growth curves. Opposed to the idea of constancy is the view that rate of mental growth is not uniform but decreases regularly with advancing years. It may be that the final outcome of the controversy depends on the further development of mental measures. There is a fair body of support to the idea of constancy and if the outcome is in TERMAN'S favour, as seems possible, it will be so much the better for practical/

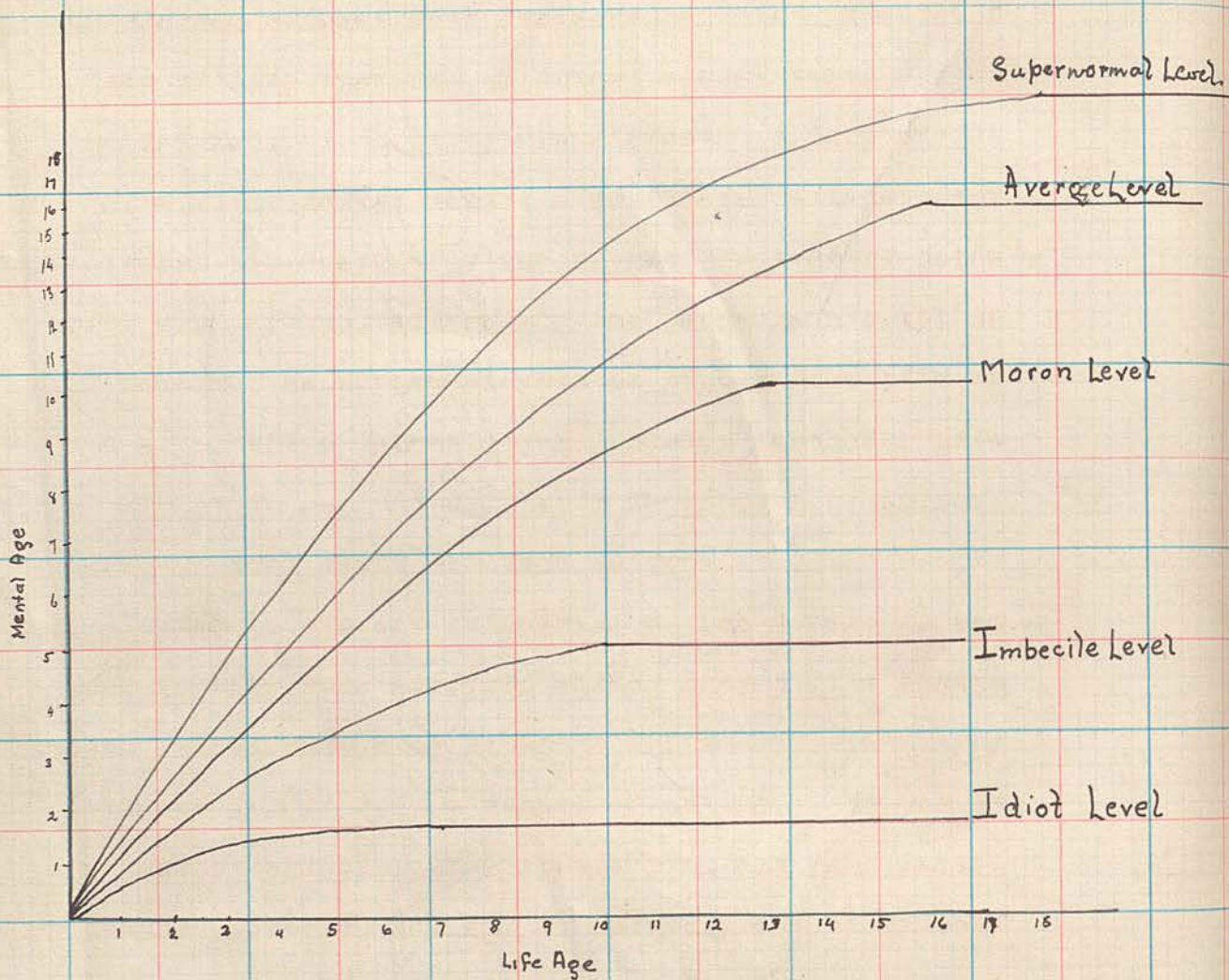
practical educational purposes.

As regards the combined operation of the two factors, ability and rate of growth, much variety of opinion exists. Whether the lower "levels" of mentality develop at a slower rate than the higher, whether all develop at a similar rate but the lower for a shorter period than the higher remains to be discovered absolutely. It is true that, with regard to the lowest mental "levels" a uniformity of opinion is generally found and it is conceded that idiots develop slowly and during a short period of time. It is as though the effort which is required to bring out the small 'potentiality' they possess readily exhausts itself. With imbeciles too, it is generally seen that rate of development is very slow and that the years of growth are few. The question of age of arrest is easily settled in such cases. But But immediately the so-called border line of imbecility is passed a variety of views is apparent. The complications are not strange. If ability were the sole factor to be taken into consideration it would be simple to conclude that the lesser intelligences reached their final goal before the greater. For, to use an expression of BURT "like a shell projected with/

with an inadequate charge, their momentum is exhausted half-way to the target". But the factor of rate comes in also as has been noted above. There are those who suggest that the moron, the average subject and the supernormal individual develop at the same rate and those who believe that they develop at different rates according to their degree of mental ability. Complications ensue. When, with some, the age of mental arrest for moronity, normality and supernormality is placed at the same chronological age, the assumption is that their differences show themselves on the mental age or ability side only and that morons, for example, have been "compensated" for lack of ability by being given a longer period in which to develop to their extreme limit. On the basis of similar reasoning the supernormals are metaphorically penalised in terms of years, for superior ability which expands with greater speed to a point of exhaustion or complete growth. That the "compensation" is paid twice over seems unjust. The feebler intelligence is given a slower rate and is also given longer time! On the other hand there is the view, supported here, that "unto him that hath shall be given and unto him that hath not shall be taken/"



# Suggested Mental Growth Curves



Scale { 3 units horiz = 1 year  
 3 units. vert = 1 mental year

taken away" - with reference to the duration of development. Altogether apart from results pertaining to the upper levels (moronity, average ability and supernormality) there seems to be no logical reason for granting imbecility a longer period of life in which to develop than idiocy to express the matter so for convenience, and not permitting the same rule to hold with regard to average growth as compared with subnormal growth. Experimental evidence is not lacking in support of this view.

It is obvious that the final forms of the mental curves and in particular the age of mental arrest will depend on the complete knowledge of the inter-relation of the three factors discussed. Here is not the place to be dogmatic, but it may be suggested that the most adequate form of the curves, leading to the fixation of the age of mental arrest for all degrees of ability seems to be such as is drawn opposite. The suggestion is made with an understanding of the manifold factors which contribute to the final result and with a recognition of the infancy stage of investigations which seek to discover the truth.

PARALLELISM/



## PARALLELISM BETWEEN MENTAL AND PHYSICAL GROWTH.

---

Recent additions to the study of mental and physical growth have contributed much to the understanding of their development from birth to maturity.

Of the two, physical growth curves are more easily established. Study extending over a large number of years such as that of Baldwin<sup>(1)</sup> verify certain general beliefs regarding growth. The author says, "In the 1914 bulletin it was discovered that the increase of growth in height is comparatively uniform for each individual, so that the growth curves enable one to prophesy with a high degree of accuracy, how tall a child of normal growth will be in a subsequent age, providing his or her relation to a given medium or norm is known. In brief, tall children do not become short; neither do short children become tall under normal conditions. This discovery has been verified again with these new data".

On the assumptions that mental growth on the average is regular from birth to maturity and is measurable at least to some extent by the scales in use, the attempt is being made to establish curves of/

---

(1) BALDWIN: From the Univ. of Iowa Studies in Child Welfare. Vol.I. No.1

of mental growth which will have the same predictive value. Terman<sup>(1)</sup> writes "The problems relating to native mental abilities and to the developmental changes which come with maturity are among the central problems of educational psychology. We cannot intelligently plan a child's later education without means of forecasting what abilities he will possess at a given time in the future. Anything that adds to a knowledge of mental growth is of practical significance for education. It is safe to predict that I.Q. 75 will never attain I.Q. 125 or vice versa. These predictions are, of course, very rough".

Baldwin<sup>(2)</sup> has presented data confirming the similarity between mental and physical growth curves. But, in indicating his belief, acceptable or otherwise, that the average curve of superior children presents a different level of development from that of average children and that the curves become increasingly dissimilar with age, he at once suggests the greater complications arising on the matter of mental curve constructions which have been indicated/

- 
- (1) Terman: Mental Growth and the I.Q. Journal of Educ. Psychology. Sept. 1921.  
 (2) Baldwin: From the Univ. of Iowa Studies in Child Welfare. Vol. II. No. 1.

indicated in the previous section.

On the subject of the rate of physical growth in children BOAS<sup>(1)</sup> writes "During adolescence the rate of growth is considerably accelerated and decreases again after physical maturity has been reached. There is a maximum in the curve before sexual maturity and after that soon zero growth. The acceleration is sooner with girls".

Several studies have been concerned with detecting whether such an "adolescent spurt" as BOAS shows to exist in the physical side, is also characteristic of the mental growth preceding the supposed upper limit of growth. TERMAN has refused to accept *this* — ? view. He writes "The much talked-of adolescent spurt looks like a myth". BROOKES<sup>(2)</sup> in the study of 171 children retested for three years found the rate of growth constant and no evidence of an adolescent spurt. He admits individual irregularities in mental growth which can only be discovered by repeatedly retesting the same individuals over a long period of time.

BALDWIN/

- 
- (1) BOAS: The Growth of Children. Science Vol.36  
 (2) BROOKES: Change in Mental Traits with Years.  
 reviewed by Parker, Journal of Educ.  
 Psychology 1922.

BALDWIN & STECKER<sup>(1)</sup>, however, in the second work previously indicated, found that there is a significant change in the trend of the curve on approaching adolescence and that an adolescent superiority of girls is also revealed. Further he maintained that the prepubertal increase appears earlier in superior children.

MURDOCH & SULLIVAN<sup>(2)</sup>, comparing the height, weight and I.Qs of 580 children of each sex from 6-18 years found some evidence of an adolescent increase in the rate of growth.

The conclusions from these studies are conflicting. It appears that the parallelism, if it exists in this respect, remains yet to be fully revealed.

An interesting study of a different nature, by PORTEOUS<sup>(3)</sup>, endeavouring to disclose a parallelism between brain capacity and growth of intelligence may be mentioned by reason of its bearing on the central/

- 
- (1) BALDWIN & STECKER: Psychological Bulletin Report 1923.  
 (2) MURDOCH & SULLIVAN: Ibid; also Journal of Educ. Psych. 1922.  
 (3) PORTEOUS: Temperament and Mentality in Maturity, Sex and Race. Journal of Applied Psychology. March 1924.



central topic of the research. PORTEOUS contends that his extensive studies of brain capacity, as derived from three different head measurements, contradict the findings of the psychologists regarding the limits of mental growth. "So far from brain growth being complete at 16 years, in the male, it will be seen that it continues right on to a period between 20 and 30 years of age. If mental growth ceases at 16 years of age then the additional brain growth has no connection with the growth of intelligence. This is an unlikely assumption". In order to establish a connection, if any existed, between brain capacity and intelligence, a mental test was given to 200 cases at either extreme of ability. From the results he concludes "The relation to intelligence is obscured because there are both large-headed and small-headed, but there are four times as many (feeble-minded) among small-headed". This only applies to the group and not to individual cases. Unfortunately for PORTEOUS' belief in the parallelism between mental ability and brain capacity, the results do not throw much light on the matter, since they are inadequate for decisive purposes. Until more direct relationship between brain capacity and intelligence is/



2  
c  
is proved, PORTEOUS' opinion regarding the continuance of growth of intelligence to 25 years, must be relegated to the background, and a warning issued against extending the 'parallelism' notion too far. Further, the contention that, on the average the development of brain capacity of women comes to a stop before that of men remains to be supported on the side of intelligence by discoveries not yet made in the employment of mental tests.

## HISTORICAL BEARINGS ON THE QUESTION OF MENTAL ARREST.

---

The problem of the age of mental arrest, it has already been seen, does not easily separate itself from adjacent or complementary problems of mental growth. Hence the review of historical contributions must include all aspects of the situation. The age of mental arrest depends on initial ability, on rate of growth and on duration of growth. It is convenient to review the matter from the side of differences in ability and to observe what different writers have to say with regard to the mental growth of idiots, imbeciles, morons, normal and supernormal subjects.

### NORMAL OR AVERAGE GROWTH.

---

. In the previous exposition regarding the nature of the controversy, certain opinions bearing on the question of mental arrest were stated. It was seen that Terman placed the age of arrest at fifteen years (most recently) and that in this he was supported by a number of investigators. The Army results pointing to a thirteen-year upper limit afforded strong opposition to this view and a tendency was evident/

evident for some investigators to avoid these extreme placings and to fix the average upper limit somewhere between the two ages. The 16 year old upper limit was derived from the practical application of the BINET-SIMON scale to adolescents and adults. Of all the general forms of the Scale Terman's alone extended far enough to fringe the mental arrest problem in its relation to average subjects.

BINET, the originator of the scale, was chiefly concerned with problems of mental defect. His scale does not extend far enough to indicate directly or sufficiently, a cessation of mental growth in normal subjects. There are those who suggest that BINET'S difficulty of supplying tests for the later age groups was due to the slowing down of mental growth at about 12 years. This may be true. The difficulty was certainly there. By the construction of the scale the rate of normal development is unity and is assumed to be constant for all ages.

GODDARD, like BINET confined himself largely to an extensive study of feeblemindedness, but his scale was standardised on the bases of an examination of very many school children. The apparent/

apparent tendency of GODDARD'S median mental age to fall off after 10 years is not a true indication of cessation of mental growth, but an unfortunate result of the testing of unsuitable subjects at the upper ages. Regarding the level of mental arrest for normal subjects. GODDARD writes <sup>(1)</sup> "The indications are that normal development ceases as a rule, somewhere near the completion of the adolescent period".

<sup>(2)</sup> KUHLMANN expresses his belief thus "The development of intelligence stops approximately at the age of 15, so far as it has been determined by these tests, and the scale (his own) registers mental ages up to this point". A mental age ought not, on this account, to be divided by a figure larger than 15, when the I.Q. is required. Subjects over the average adult level, cannot be given a mental age.

<sup>(3)</sup> More recently BALLARD has discussed the actual problem. The curves of mental growth which

ne/

- 
- (1). GODDARD: Human Efficiency and Levels of Intelligence. 1920.
  - (2). KUHLMANN: Handbook of Mental Tests.
  - (3). BALLARD: Group Tests of Intelligence: Chapter on the Upper Limit of the Growth of Intelligence.

he quotes from WOODROW, are drawn on a later page. Suffice to say here that he presupposes the same 'life age' of arrest for subnormal, normal and supernormal growth, and fixes the age at 16 years. His conclusion is based on the use of his Absurdity Tests with children of all ages. Results from these tests indicated a slowing down of mental growth after twelve years of age, with normal children. BALLARD comments on the findings of PORTEOUS regarding the mental growth of Australian aborigines, which pointed to an age of mental arrest several years sooner than that for white children. A discussion of BALLARD'S reasoning regarding the age of mental arrest is reserved for later comments.

#### SUBNORMAL GROWTH.

The literature on this side is more plentiful and worthy of consideration for the light it throws indirectly on the question of normal growth arrest.

BINET'S interest in the problems of subnormality was profound. His studies of normal mental growth were often intended to throw light on mental defect. He writes "Understanding the normal progress/



progress of intellectual development, we shall be able to determine how many years an individual is advanced or retarded, in a word to determine to what degrees of the Scale idiocy, imbecility and moronity correspond". He desires <sup>(1)</sup> further to know if "states of intellectual inferiority are caused by arrested development or by very slow evolution continued regularly or intermittently, or to see if some essential factors would increase while others remain stationary or undeveloped". The impossibility of predicting future development from a single mental age determinant is noted when he remarks "It is understood that these diagnoses apply only to the present moment. He who is an idiot to-day may with the progress of age become a moron; or on the contrary he may remain an "idiot for life". One knows nothing of this, the prognosis is reserved". The upper limit for moronity BINET places at about 12 years.

(2)

GODDARD writing recently states that there is an efficiency limit for each individual which, when once attained, cannot be surpassed. "The study of feeble-mindedness has confirmed our own belief/

- 
- (1). BINET: The Development of Intelligence P.143  
(Kite). L'Année Psychologique. 1908
- (2). GODDARD: Human Efficiency and Levels of Intelligence.

belief that intelligence is a matter of brain cells and neural patterns, and, still more definitely it is a question of the development of the larger association areas of the brain, the function of which develops relatively late and hence the development is particularly liable to arrest. Moreover, when arrest has taken place there is no evidence that it ever starts again. Experimentation has taught us that if the mental level is as much as three years lower than the chronological age, it is practically safe to assume that arrest has actually taken place and that the level will never be higher or at least significantly so. In another place <sup>(1)</sup> GODDARD makes the statement "Feeble-mindedness is clearly an arrest of development at some time previous to adolescence - a stopping of mental development all along the line. The arrest is not necessarily sudden. There is usually slowing down covering several years". A qualification of this opinion is expressed thus <sup>(2)</sup> "How does the child reach his point of arrested development? Is it by a gradual slowing down or by a sudden stop? ..... Perhaps it is a matter/

---

(1). GODDARD. Feeble-mindedness, its Causes and Consequences. 1914. P. 546.

(2). DOLL: P. 19.

matter of the individual. Some develop one way and some another.....There are some indications, however, that some children are pretty nearly normal until about eight or nine or ten years of age, and then they drop rather suddenly.....Others seem to have always been a little backward, but keep getting more and more backward, until the age perhaps of nine or ten, when they seem to have entirely stopped". And conclusively GODDARD writes<sup>(1)</sup> "The slowing down process may begin at any time during childhood. Of course, with the idiot the matter is relatively quickly settled; with the imbecile less quickly and with the moron it becomes a problem for a number of years.....It does not seem likely that all such defectives have the same rate of slowing down; some probably reach the stopping place more quickly than others, just as some begin the slowing down process at an earlier age than others".

As regards the age of arrest on the mental side, GODDARD designates as idiots, those having a mental age of one or two years. Imbeciles may be classified/

---

(1) . Psychology of the Normal and Subnormal.. P.16.  
quoted by DOLL P.19.

classified as Low Grade, Middle Grade and High Grade and are between three and seven years mentally. Morons have a mental age between eight and twelve. This classification was confirmed by CHOTZEN.

BOBERTAG's (1) experiments lead him to conclude, "Children whose intelligence is above average gain more rapidly than those whose intelligence is average while those below average gain more slowly than those who are average. Also within the group of those who are average it must come about that some who are average must become retarded and some become advanced (in units of the difference between the actual mental age and the standard mental age for the life age of the subject.) So, BOBERTAG prefers to use I.Q. ratio in expressing feeble-mindedness or any deviation from the normal. But the constancy of the I.Q. he cannot accept and he rejects the form of curves which the constancy of the I.Q. implies. (P. 532) and which he expresses thus:-

"Das Entwicklungs tempo der Kindlichen Intelligenz gleiche der Aufwärtsbewegung eines Lifts. Das begabteste Kind entsprache dem am schnellsten fahrenden Lift,/"

---

(1). BOBERTAG: Zeitschrift für Angewandte Psychologie 1912. P. 528.



Lift, das unbegabteste dem am langsamsten fahrenden". The curves of development are wrongly drawn when they have a straight line form (GERADE-KURVE); they ought rather to have a "projectile" form (WURF BAHN-KURVE). BOBERTAG has several reasons for his rejection of the curve forms, not the least among which is that CHOTZEN'S results with the Hilfsschule children indicated that mental decrease was not absolute from year to year but increased relatively with increase in age.

(1).

KUHLMANN is reported by DOLL to believe that feeble-minded growth continues up to 15 years for any degree of defect except idiocy. He is not consistent in reference to rate of development, in one place saying that it is constant and challenging BOBERTAG'S view that it is a decreasing variable, at another time supporting the contrary belief. More recently (2). KUHLMANN concludes that mentally defective children develop at a rate proportional to the grade of intelligence and that the rate of development decreases with age for all grades. His conclusions/

- 
- (1). DOLL: The Growth of Intelligence P. 95.  
 (2). KUHLMANN: Journal of Applied Psychology.  
 Sept. 1921.



conclusions may be summarised as follows:-

- (1) The Mental ages of feeble-minded increase with age and at a rate proportionate to the degree of mental efficiency.
- (2) On the whole the mental age ceases to increase between the age of 15 and 18, for idiots about three years earlier than for borderline.
- (3) The lower grades lose more frequently in mental age than the higher grades.
- (4) The frequency in loss in mental age increases with age independently of grade.

Regarding the rate of mental development  
 (1). STERN remarks " Since feeble-mindedness is a condition of development that is below the normal condition, the rate of development will also be a slower one and thus every added year must magnify the difference in question". The absolute difference between mental and chronological age is not satisfactory for expressing feeble-mindedness, Since a year of mental growth means a different thing at different stages. He therefore uses the I.Q. and of it says "The quotient does not seem to afford an actually constant expression of the degree of feeble-mindedness but shows a tendency to fall in value as age increases. This tendency, it is evident, is but slight/

- 
- (1). STERN: Psychological Methods of Testing Intelligence. P. 79.

slight within the limits of age that have been mentioned, and so for many problems it may be disregarded. Before and after these ages the fall in the value seems to take place more rapidly. In the case of the later age levels this is easily intelligible, for once the age of arrest is reached the quotient obtained by dividing the mental by the chronological age must decrease as chronological age increases". In other words the I.Q. can only hold during the period of mental growth.

Then later (P.84.) he writes:- "The feeble-minded child, it must be remembered, not only has a slower rate of development than the normal child, but also reaches an age of arrest at an age when the normal child's intelligence is still pushing forward in development".

In a later<sup>(2)</sup> writing in speaking of the mental growth curves STERN emphasises the difference in age of arrest for normal and subnormal children in pointing out that both normals and subnormals lessen in development as they near the end of their intellectual growth period but with the subnormals the/

---

(2) Zeitschrift für Angewandte Psychologie. P.4.

the time comes sooner. His own words are:- Zwei mit abnehmender Geschwindigkeit Steigende Kurve können doch sehr wohl in einem konstanten Verhältnis zueinander stehen, indem die Ordinaten der niederen (Schwachsinnss Kurve) sich durch einen konstanten Faktor z.B.  $\frac{1}{2}$ , von den Ordinaten der höheren Kurve unterscheiden. Dieser entsprechende Verlauf muss allerdings dort ein Ende haben, wo die Entwicklung des einen aufhört, während die des anderen noch weit geht". In the same passage STERN takes exception to the conclusions of BOBERTAG that the development of mentally defectives progresses at an ever decreasing rate on the grounds that the subjects tested were unsuitable for proving or disproving the constancy of the I.Q. The same criticism is applied to the other writer's results.

(1)

TERMAN finds that mental growth develops at a constant rate for all degrees of brightness and dullness, except idiots and low-grade feeble-minded and that "the mental age of high-grade morons appears to change little after the age of 14 or 15 years".

BURT/

---

(1) TERMAN: The Intelligence of School Children  
The Stanford Revision of the BINET  
Simon Scale. Ch.III.

(1)

BURT observes that the predictive value of the I.Q. is not yet established. With feeble-minded subjects there is a tendency for the ratio of mental to chronological age to decrease in value with successive years. "It is small; but it is steady". But in the re-examination of 72 backward children, he found that the I.Q. remained fairly constant during the period of growth. In the upper age groups the decrease was noteworthy but BURT concluded that "such a gradual subsidence of progress must be a universal characteristic. The lower grades have no monopoly. Older defectives of all levels exhibit a premature loss of development "like a shell projected with an inadequate charge, their momentum is exhausted half-way to the target". In the same connection he adds "The progress of the typically feeble-minded <sup>waves</sup> appreciably as the end of their school career is approached: their years of growth are few and slow".

DOLL/

---

(1) BURT: Mental and Scholastic Tests.

DOLL.

---

It has seemed desirable to omit from the various parts of the review any relevant considerations that arise from the work of DOLL<sup>(1)</sup> who, in support of the American Army results contributes one of the few existing books on the Growth of Intelligence. The omission was made, partly on account of the desirability of discussing DOLL'S curves of growth as against other possible curves and partly because of the contradictions in his whole exposition.

In a general introduction, DOLL exposes, sometimes erroneously, the views of the other and earlier investigators on the question of growth of intelligence. Regarding normal growth, he discards TERMAN'S views, as before stated in favour of the Army Results. His arguments against TERMAN have already been shown, but at the outset it must be noted that they refer to the 16 year old level, which TERMAN has now discarded for the year below. The opposition of DOLL challenges the standardisation of/

---

DOLL: The Growth of Intelligence 1920.



of the Stanford Revision itself. In the latter part of the book the constancy of the I.Q. is also refuted.

On the subject of subnormal growth. DOLL is more justified in speaking, since his original research is based entirely on data available from the retesting of subnormal children and adults from 1910 to 1919 at Vineland School N.J.

The value of DOLL'S investigation is lessened by the small number of suitable subjects tested. Although some 203 individual record sheets (from individuals aged 6 to 46 years, plus one aged 66) were utilised in the report only 95 of the subjects were under 15 years of age at the time of first testing and, therefore, the test had presumably already reached their final limit of growth. A further unfortunate circumstance was the employment of the original translation of the BINET SIMON 1908 Scale, during 1910 and the employment, subsequent to that date of the GODDARD or VINELAND REVISION of the Scale. This necessitated the transcription of the 1910 results in terms of the GODDARD REVISION, if they were to be preserved, moreover, it was not always possible to do this on account of the omission of/

of certain BINET Tests from the GODDARD REVISION as well as on account of modification of scoring or procedure in giving the tests. The transcribed records being often incomplete, DOLL devised a series of objective rules for scoring incomplete records which were not so incomplete as to demand rejection. A still more serious objection to the incomplete records resulted from the narrow range of the testing method employed in the early work. Another series of objective rules for overcoming this incompleteness was devised.

DOLL himself was entirely in charge of the direction and analysis of the examinations and re-examinations from 1913 onwards, except such as were carried on at Summer Schools, which were conducted by thoroughly competent individuals. Inaccuracies arising from the personal equation of ~~examiners~~ are considered to be insignificant in regard to the examination of mentally defectives.

(It seems impossible that so many limitations to the value of a work could present themselves without some at least being true. Apart from the objectivity of the method of completing the record sheets, it is undesirable that it should have been necessary at all, and it is doubtful if a series of seven/

seven rules, with a varying number of subheadings under each, attains anything like the possible accuracy of a true and adequate examination).

The record sheets completed in accordance with the rules, the mental age was then calculated. An attempt was made to standardise the time interval between examinations so as to reduce the work of the calculation. For this purpose, as well as to summarise minor fluctuations which arise in mental ages it seemed best to 'smooth' the growth records. This was done by calculating two-years averages of the successive mental ages. DOLL remarks "This procedure had the additional advantage of furnishing age intervals of a single year, which gives a convenient basis for further calculations. Each mental age was, accordingly, used twice to give the 'true' mental age. Where gaps of one or two years came in the test records, the method was to interpolate hypothetical mental ages by means of averages from the mental ages before and after the gap. As these gaps occurred most frequently with subjects over 15 years the matter was not strictly important since these subjects at least had ceased to develop. The exact life age on which the mental ages were calculated, was not ascertained so that a slight error enters into the tabulations/

tabulations on that account. In DOLL'S opinion it is justified by its simplification of the presentation of the data.

The classification of subjects is on the bases of final mental age. The desirability of classifying by life age had to be put aside because of the "small and unequal members at each age and the variations in mental ages at each life age". In the same way it was impossible to make use of the I.Q. to ~~eliminate one~~ of the two factors of age and mental age, since it ~~involves~~ two assumptions which form the very basis of the study. These are

- (a) That the level of arrest is independent of the age of arrest (the level being reached about 16 years of age.)
- (b) That the rate of development is constant for all types as well as all degrees of intelligence.

(What these remarks may mean, it is impossible to say. DOLL has not noted that TERMAN places the age of arrest for moronity at about 14 years while placing the age of arrest for normal subjects at 16. The truth is that he has managed so to intermingle three ideas, if they might be called so, viz: TERMAN, I.Q. and 16 years that the thought of the one invariably associates itself with the thought of/

of the others. Amongst other things, too, DOLL has forgotten for the moment that BOBERTAG, (whose influence on him is marked) CHOTZEN and KUHLMANN, at least, all employed the I.Q. without believing in its constancy or once hearing of the 16 year old level.) To pass on, the classification reveals the amount of increase for each subject after one year, two years and so on. The discovery is soon made that the annual rate of increase, expressed in percentage form was very small, in general.

The analysis of the data is undertaken for two main purposes.

- (1) To determine the relation between the age of arrest and the final mental age level.
- (2) To determine the relation between the rate of mental increase and final mental age.

Relation of Age Arrest to Level/Arrest.

A Limitation to analysis is imposed by the small numbers at the younger age periods. But DOLL'S conclusions are as follows:-

- (1) Substantial increases are limited to subjects who are under 15 years of age.
- (2) Only two subjects show decrease in mental age amounting to as much as one year.
- (3) Only 39 subjects gain as much as at least one year in at least 5 years. All but one are under fifteen years of age at the first a the first record i.e. 61% of the subjects under/



under 15 fail to show mental increase amounting to as much as one year in five years. Further, the lower the mental level the greater the number of subjects who have reached their final level. So, the lower the mental level the earlier the age of arrest.

- (4) It is difficult to calculate the age of arrest for each mental level on account of the small number of records from young subjects.

(DOLL quotes a table here to give some indication of the age of arrest)

Mental Age Level	1	2	3	4	5	6	7	8	9	10
Age of Arrest	10	8	6	11	12	15	15	12	15	14

- (5) There is observed among subjects who are older than the age of arrest for their mental level a slight tendency to decrease in successive examinations independently of age. (DOLL takes this to be evidence against 'coaching' in the tests and as evidence against the power of experience to affect the mental level. There is some suggestion too that institutional regime may affect individuals so as to deprive them of everyday knowledge such as the date etc.

The general conclusions are four.

- (1) That the age of arrest is associated with the level of arrest.
- (2) Within a particular final mental age group the younger subjects are more likely to improve than the older, independently of actual retardation.
- (3) Gains are restricted to subjects under 15 years.
- (4) Variability is very great and prognosis on the basis of mental age alone is practically impossible for individuals.

Relation/

Relation between Rate of Mental Increase  
and final Mental Age.

From the data DOLL concludes that most of the subjects under 15 are developing at a very slow rate. Rate of growth may be a function of life age or of mental age or of degree of retardation, apart from external influences. Individual differences being so great resort to averages is necessary. Because the validity of the I.Q. is no question, it cannot be employed in the study of degree of retardation. Correlation methods would be best for dealing with the problem but the numbers are too small.

Trial calculations indicate that the rate of growth is practically independent of mental age. Therefore it remains to be seen whether life age and rate of growth are related. DOLL'S own words are better quoted here: (P.62) "We are concerned with the amounts of increase as well as the strength of the tendency. Therefore it seemed advisable to calculate the actual average rates of increase for the successive life ages. We could then plot the average functional relation between rate and life age. If there were at any life age a correlation between rate and mental age it would be necessary to allow/

allow for this influence. But if the distribution of mental ages at each life age period is approximately the same from age to age, then an average rate of growth by life age may be established independently of mental age. Or if there is no significant correlation between mental age and rate of growth at each life age grouping, then again the rate of growth will appear as a function of life age independently of mental age."

DOLL presents a table showing the annual rate of increase for all subjects under 15 years at each life age, & relates from the table that there is no consistent dependence of rate of growth upon mental age at any life age grouping, on the other hand, rate of growth seems to be a direct function of life age. The obvious conclusion is that "the average rate of mental age increase of these feeble-minded subjects decrease regularly and steadily between the ages of 8 and 14 years and reached a practical <sup>minimum</sup> ~~maximum~~ at 13 or 14 years." Lest this conclusion seems to be in contradiction to the previous conclusion that the age of arrest is a function of the level of arrest the data from improving subjects are studied alone. The results of the 38 improving subjects/

subjects confirm the previous findings and point to a slight negative correlation between rate and degree of retardation. DOLL then proceeds to study the relation between age of arrest and rate of growth. He plots the average growth curves for each final mental age group as well as the limited numbers permit. From the curves he is able to state more definitely than before the relation of mental age to age at arrest.

Final Mental Age.	1	2	3	4	5	6	7	8	9	10
Age of Arrest	?	?	7	11	11	12	12	12	15	15

"In terms of the grades of feeble-mindedness we may say that the idiot and low-grade imbeciles are arrested probably below 7 years of age, the upper-grade imbecile about 12 years and morons about 15 years"

The average rates of growth are practically negligible for final mental ages 1 year to 4 years inclusive after the life age of 6 to 10 years. For final mental age groups 5 to 9 the growth rates are approximately the same between the life ages of 8 and 14 years.....At all the mental ages where the subjects are young enough to furnish data for the rate of growth the rate is a progressively decreasing variable./

variable. Moreover, inspection of Fig.10. lends colour to the theory that the upper grades of feeble-mindedness have approximately an average normal rate of growth early in life."

This concludes DOLL'S original investigation. The results support the initial beliefs as expressed in the words - "Logical considerations lead to the belief that the age of arrest and the rate of retardation are both functions of the level of arrest, on the average at least. Psychological and logical considerations indicate that for the feebleminded it is probable that the rate of growth is a decreasing variable".

In addition to the original data, DOLL makes use of some material given by Miss Gillingham, whose interest is in the study of supernormal children. The data bear on the test results of 35 superior children of life ages 9 to 12 years at the time of first examination.

Regarding the age of arrest for supernormal children DOLL remarks that "it is safe to assume that it is not earlier than 13 years of age. The rate of growth of superior children will depend partly on life age, and also partly on the mental level/



level. As a postulate DOLL suggests that to counterbalance the decreasing rate tendency of the subnormals the supernormal rate may be an increasing variable. From the actual data at hand definite conclusions cannot be made. The numbers are too small at each separate age. But DOLL observes "We may say that in the case of children of superior intelligence between the ages of 8 and 12 years inclusive, the rate of development is apparently independent of age, thus, it is fairly constant, except for the influence of mental age. There is only a faint and uncertain tendency for the rate to increase with age. The rate is, however, somewhat related to mental age during these periods".

This/

This then is the sum of DOLL'S effort to obtain the upper limit of the growth of intelligence and to study mental growth generally by the method of individual "BINET" testing. It is time now to turn to his theoretical curves of growth to see how they are borne out by the results of the investigation.

DOLL'S initial theoretical considerations of the problem and the various contributions towards it led him to construct curves of growth after the manner of those drawn opposite. Curves representing the growth, taken on the average, of idiots, imbeciles, morons and normal subjects are included. Of a dissimilar type are those suggested by BALLARD<sup>(1)</sup> and, whereas DOLL supports the Army results and a 13 year old limit of intelligence growth, BALLARD adheres to the 16 year old level. The form of the curves which DOLL outlines is in keeping with his thesis and with the results which he has reported from the testing of many feeble minded children and adults.

The average idiot progresses during the period of growth at a very slow rate and reaches a final age of arrest at about 7 years life age. The imbecile/

---

(1) BALLARD: Group Tests of Intelligence.

imbecile with somewhat more average celerity to a final level of 9 or 10 years (DOLL'S actual results say 12 years). The average moron curve still

"sleeper" in the growth period shows a gradual decrease in rate and a flattening which continues to the age of arrest at approximately  $13\frac{1}{2}$  years.

Lastly, the adult curve, presumably conforming to the principle of the BINET Scale Construction, where a year of mental age corresponds on the average to a year of life age, manifests a still greater average rate of development, but, an arrest of development at approximately the same age as is attained by the average moron. Referring to the subnormal level of arrest in the course of his argument DOLL says:-

"It is probably something more than a coincidence that this age of arrest of the average rate of feeble-minded growth coincides with the age of arrest of average normal subjects previously discussed".

Probably; but it would have been more convincing if a reason had been attached to the statement.

Taken at their face value,  
DOLL'S curves present a glaring contradiction/

sharp bend in the theoretical curve at 13 years so that thereafter there is zero growth immediately indicated. Now, in reality, no growth curve or any other curve ever did behave in this fashion. Even supposing that he has drawn the curve purposely after this fashion he at once reveals a weakness of the plan. There is no indication of the 'slowing-down period' of growth that precedes mental arrest and as the curve is drawn it could not easily be indicated so as to keep the mental age at 13 without displacing the chronological age of arrest at least two years. This is, however, pure surmise and must not be considered overmuch.

Regarding the subnormal age of arrest and rate of development DOLL has claimed to have found results which fit in with his theoretical curves. His results, however, have to stand the test of severe criticism on the ground of the inadequate number of suitable subjects tested and the smoothing process adopted in the treatment of results which as TERMAN<sup>(1)</sup> points out inevitably flattens the curves/

---

(1) TERMAN: Mental Growth and the I.Q. Journal of Educ. Psych. 1921.



curves of growth.

Supposing for the moment that DOLL'S position stands unchallenged by any results, it may yet be asked if it is theoretically sound. Both the curves of DOLL and of BALLARD establish an average upper limit for the growth of subnormals, average subjects and supernormals at the same life age but at different mental ages.

BALLARD,<sup>(1)</sup> writing on the upper limit of average intelligence and the possibility of fixing it at 16 years, writes: "If after 16 years of age the curve of growth is virtually horizontal, as it seems to be, we are forced to conclude that this is not the average age at which growth ceases, but the maximum age. For the general curve is the resultant of a large number of individual curves and if the resultant curve is horizontal after 16, then, either all individual curves are horizontal after 16, or if some rise others must fall. In other words, we must choose between two alternative suppositions, one that nobody improves after 16, the other that some people in sound health begin at that age to suffer from senile decay". If this view is not based on the idea of subnormal average and supernormal growth/

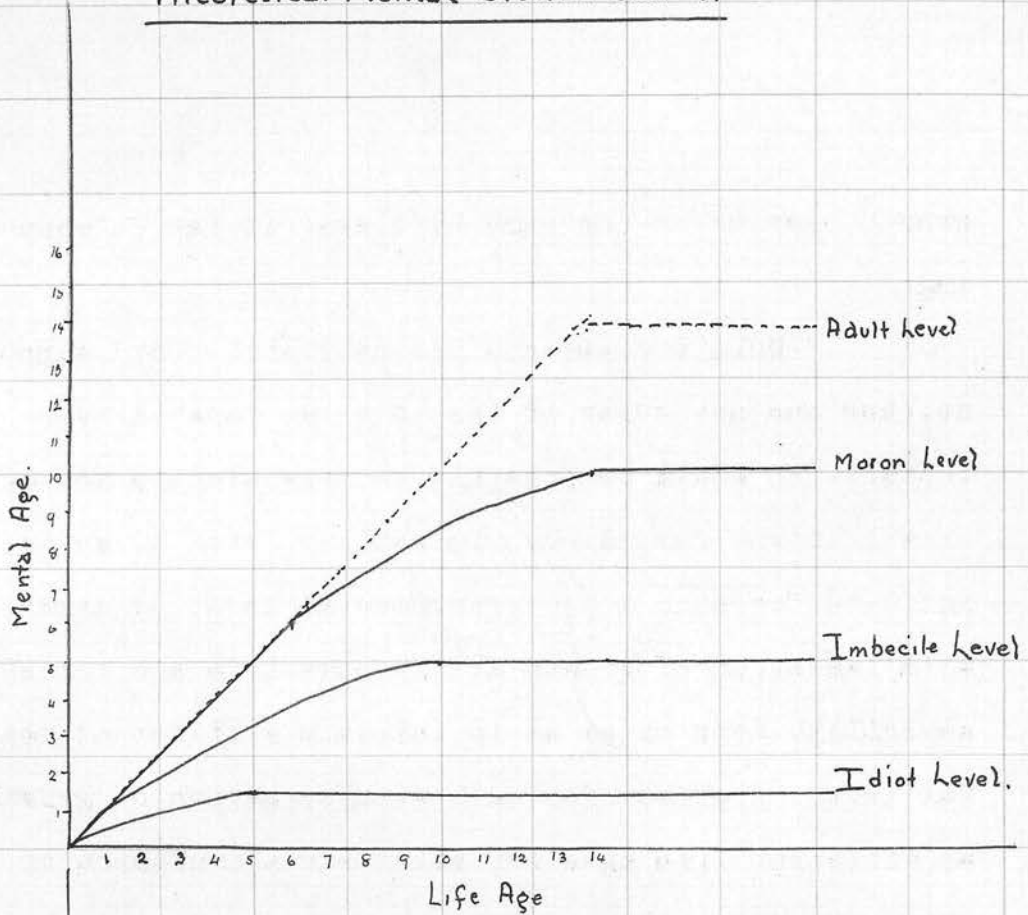
---

(1) BALLARD: Group Tests of Intelligence. Chapter on The Limits of Intelligence.

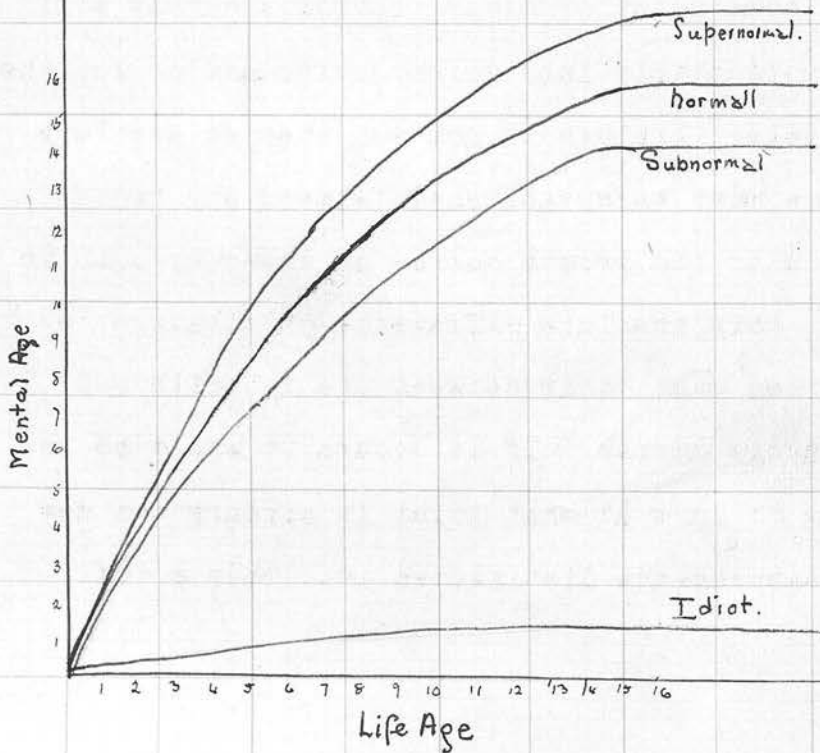


# Theoretical Mental Growth Curves.

Doll



Ballard



growth ceasing at the same life age, it has no meaning.

DOLL's views are inconsistent. They support neither one nor other of the only two tenable hypotheses. It would be possible to draw average mental growth curves for idiocy, imbecility, etc. so as to indicate for each a different mental level of arrest with cessation of growth at the same life age for all, as BALLARD does or so as to indicate a different mental level of arrest for each with cessation of growth at different life ages for all. But a confusion of these two principles is not permissible. It is obvious that the average growth curves indicated are the resultant of many individual growth curves. It follows that, at some point or other, imbecile curves will shade imperceptibly into moron curves and so for the other levels. If this is not so, then an absolute difference must be established between any two levels. With the growth curves as drawn by DOLL in his book, this absolute difference or cleavage between curves must occur between the imbecile and moron average curves. If it occurs it would be interesting to know at what point it appears and how mental measurements distinguish it. This point evades/

evades attention when it is forgotten that the average curves (regarding whose artificiality comment has already been made) are only a convenient method of expressing the individual growth curves. Moreover, the scale that measures ability is a quantitative scale and 'quality' is only implicit. On this account, it would not suffice to attempt to explain the contradictory structure of the curves as ~~BALLARD~~<sup>60</sup> and DOLL draw them by anything outside the possibilities of mental testing or beyond its realm. This point serves only to emphasise the inconsistency.

DOLL'S curves, or in particular, the form of the normal curve which he draws, is not supported by his theoretical arguments as far as these extend. If "logical considerations" lead him to believe that "the age of mental arrest is a function of the level of arrest", they lead others to believe the same. Summarising his conclusion, he states:- (P 4a)

"The lower the mental level the earlier the age of arrest" and proceeds to add tables, already given above, in support of his argument. But the next "logical" step surely is to suggest that, since idiots are arrested before imbeciles, as he has proved and others knew before, and since imbeciles are arrested before morons, as ~~one~~ <sup>one</sup> would have believed, then morons are/

are arrested before normal subjects on the average. In argument DOLL supports what in theory he will not believe.

Added to all this are the facts concerning the collection and treatment of the data. The full outline of his work has been given to show the manifold unfortunate circumstances that detract from its worth. The value of the test sheets from the early investigation is questionable. The smoothing of the data has its advantages, no doubt, but flattens the growth curves for each individual. Sundry minor criticisms might follow regarding inconsistent and false statements in the report but the other aspect of the problem merits more attention. As regards the problem of rate of subnormal development, DOLL is no more fortunate than in his considerations of the age of arrest.

At the outset "logical considerations" and psychological considerations lead DOLL to suggest that "for feeble-minded it is possible that the rate of growth is a decreasing variable". The number of subjects tested considerably hinders the solution to the question.

After grouping the results under the different life ages and arranging records within each life/

life age in order of mental age, it is true that he finds no general correlation between rate of growth and mental age. This leads him to emphasise that "rate of growth is practically independent of mental age". (P.62) By comparing the average annual rate of growth for each life age he concludes that "it is instead a function of life age". It is desirable that DOLL should get this result, to support his thesis that constant rate of growth is not characteristic of subnormality. But common sense rebels against his argument and seeks to defend itself by a thorough analysis of the data. He tries to show that within each life age group with the individuals ranged in order of mental age, there is no evidence of greater rate of development in the case of those having higher mental ages than in the case of those having lower mental ages. What then is the significance of such a statement as this, used by BURT<sup>(1)</sup> "The average progress made by the entire group (of subnormals) during the year was +0.47 mental years, a rate of barely one-half a mental year per annum" and what can be said of it as against another and suggestive/

---

(1) BURT: Mental and Scholastic Tests. P.155.



suggestive one of DOLL'S statements: "the rate of growth is inconsiderable in the lower mental ages"? The suggestion in the latter case seems to be that the rate of growth is not so inconsiderable in the higher mental ages. Then this is a coincidence and no more!

There is a qualification to DOLL'S conclusion incidentally remarked on by the author and it is this "There is some tendency at each age for the rate to increase with mental age from mental age one year to about mental age 5 years or 6 years; after the point the rate tends to decrease. The net result of these compensating tendencies is to produce a negligible correlation between rate and mental age at each life age".

The evidence in support of the dependence of rate of growth upon life age might prove DOLL'S thesis were it emphatically or indubitably apparent. But, as it happens, the results are neither emphatic nor indubitable.

The first tabulation from which the conclusion is derived includes all subjects under life age 15. But of these DOLL recognises that 61% are non-improving subjects, and their slight fluctuations of/

of growth in years beyond the age of arrest may obscure the true findings. Accordingly he gives another tabulation excluding the unsuitable records. The subjects who remain are only 38 in number and are divided among the life ages 7 to 14 years with one peculiar case at 19 years. Beyond 9 years, four individuals at most represent each life age; the greater representation numbers eleven. Therefore averages are not reliable. Moreover, of the 38 subjects at least half are imbeciles so that the moron results are based on less than twenty cases. This goes to show that DOLL'S moron 'rate of growth' conclusions may not be justified at all except regarding these special cases. His conclusions are as likely to be untrue as true. No doubt the number also affected the fixation of the age of arrest for morons. Chance selection seems to have resulted in it being placed too high.

If DOLL has not succeeded in finding that the mentally defective rate of growth is constant he has certainly not succeeded in proving the contrary. If he contends that his data go to show a correlation of rate of growth with life age it is possible for others not only to indicate the insufficiency/

insufficiency of his proof but to deny it altogether. The majority of the 'improving subjects' are imbeciles. DOLL'S results bear chiefly on their growth. He indicates, as before quoted "There is some tendency at each age for the rate to increase with mental age from mental age one to about mental age 5 or 6 years" and so on. That is to say, he submits that there is a dependence of rate of growth upon mental age in this period, and so disproves his own inexplicable thesis that rate of growth is independent of mental age. Beyond this period there are few subjects - too few subjects on which is based any conclusions regarding this matter with reference to the moron level.

This by no means completes the criticism that DOLL'S argument and results could receive. DOLL is not always clear and often inconsistent. His conclusions are variable and suggest some dubiety in the writer's mind regarding the beliefs he expresses. His curves of individual growth fail to support his argument concerning the relation of rate of growth and mental age, and that from inspection alone/

alone. The validity of the records and scoring have already been commented on. Terman,<sup>(1)</sup> in reply to DOLL'S refutation of the I.Q. remarks that after a GODDARD mental age 7 a year of growth as measured by GODDARD'S Revision is 1.2 years by the STANFORD-BINET Revision. The result is the production of an "exaggerated flattening of growth curves".

The general criticism reveals one thing at least, that this work, suggestive in some ways as it is, scarcely merits the lengthy discussion it has been allowed. The reason for this discussion was not solely to expose DOLL, but also to lay a convenient foundation for the outlining of another view by referring to the form of curves of growth which he designs.

#### ALTERNATIVE VIEW.

This other view has already presented itself in the course of the criticism. It is that idiots, imbeciles, morons, average subjects and superior subjects are not only distinguishable by reason of their different mental ages but also by reason of the different period of life age at which they cease to develop. The curves opposite <sup>Page</sup> express this pictorially.

DOLL'S/

---

(1) Terman: Mental Growth and the I.Q. Journal of



DOLL'S view is rejected as being inconsistent; the view that all 'levels' show signs of cessation of mental growth at the same life age is rejected because it cannot be made to apply to the lowest grades of mentality as far as they are measured by test methods<sup>+</sup> and because adequate experimentation with the lower levels augmented by DOLL'S inadequate experimentation with the lower levels, proves the contrary. On the same principle DOLL'S  $13\frac{1}{2}$  level is set aside as being nearer to the level for moron subjects, though probably too high. The experimental evidence from the work of many thorough psychologists stands directly against the Army results. The arguments launched against the Army testing are mostly indisputable though some of them are inapplicable to other situations as well.

Lastly the results of the slight investigation that follows, inadequate as it is in many respects, at least makes one thing clear that from the sample of Scottish children tested the possibility of placing the average upper limit of the growth of normal intelligence at  $13\frac{1}{2}$  years is unmistakeably small.



## INVESTIGATION.

---

### Purpose.

---

The purpose of the investigation is to determine the upper limit of the growth of intelligence so far as it can be measured by the use of mental tests. From the foregoing discussion it is evident that, so far, little definiteness has characterised the attempts to decide the question and that in fact, those investigators who have contributed to the problem at all have mostly done so rather indirectly than by design.

Of the difficulties which lie in the way of such an investigation, two are especially to be mentioned. First, there is the difficulty which the present stage of the development of the tests themselves presents. Twenty years only have elapsed since BINET made the initial attempt to measure intelligence by definite quantitative means. Most of the investigations of worth since BINET's time have been concerned with trying out and modifying his tests under every possible condition to all available subjects of every age. TERMAN's Stanford Revision, the most/

most efficient of all, ~~date~~ only a few years back BURT's thorough standardisation of the scale for use with English children is even more recent, though presenting the labour of many years. Group tests date roughly from the American Army investigation. Therefore, it cannot be supposed that the measures employed, be they of the individual or group variety, are altogether satisfactory. Perfection is the aim, but those who have laboured hardest are least satisfied to-day because the goal seems very remote. It is only possible to make one step at a time and that with much hesitation.

Another difficulty is the finding of suitable subjects. Investigations with intelligence tests have been confined largely to the two extremes, to the testing of children in the elementary schools and to the testing of adults in attendance at colleges and universities. The intermediate group has, for the most part, escaped, owing to the fact that children leave school when the age limit of compulsory attendance is reached, namely at fourteen years, and thereafter cannot be found grouped together except in evening Continuation Schools.

For the purpose of an investigation bearing/

bearing on the upper limit of the growth of intelligence this is precisely the group that it is most desirable to test. Any other group tested beyond 14 years tends to be influenced by the factors of selection, factors which exercise an increasing influence as the years pass, and thereby renders the experimentors' task of making conclusions from results an exceedingly difficult one.

In face of these two difficulties, it was necessary, first of all to determine what tests would be most suitable for the possible solution of the problem, and, in conjunction with this, to take every possible measure to obtain subjects whose test results would go to make the solution possible.

It appeared that, taking all things into consideration, it was better to resort to group testing than to employ individual test methods. The latter, if need occurred and if time permitted, could be applied later. The disadvantages of the group test method - the failure to disclose individual differences and the concealment of individual rate of growth did not seriously affect the attempt to discover the average upper limit of mental growth.

Mental/

Mental testing, historically regarded shows that, parallel with the movement towards increased exactitude in measurement has existed the aim of improving the tests employed in experimental investigation. The improvement lies in the discarding of tests by simple processes in favour of those bearing on the complex mental processes. One of the earliest experiments with tests of complex mental processes was DR. BURT. As the outcome of his work, he was able to show that higher correlations existed between these and intelligence otherwise estimated than between simple mental processes and intelligence otherwise estimated. Since this fact became generally known, intelligence test scales have embodied more of the complex types of test. Pure performance tests are often omitted on account of their lack of appeal to higher processes. Group tests in America have relied more or less on the higher mental processes. Accordingly the attempt was made here to select tests which would be most suitable for detecting ability at the upper limit of intelligence growth and disregard tests of discriminative worth only at the lower levels.

Subjects/

## SUBJECTS.

The first steps taken to find suitable subjects for testing purposes were, of necessity, limited to certain spheres. The mental tester is not welcome at all doors. Through the interest of DR. DREVER, under whose supervision the investigation was carried out, and by the courtesy of the headmasters, access to two schools was readily obtained. These were the James Clark School for junior technical education and George Watson's Boys' College.

The former recruits pupils of both sexes between the age of 12 and 15 years, from the ordinary Board Schools of Edinburgh. It offers three possible courses of study, a boys' technical course in which handwork is uppermost, a girls' technical course, which emphasises domestic science, and a commercial course leading to clerical work of various kinds, which is followed by both boys and girls. Concurrently, the general education of the pupils is kept up.

George/



George Watson's Boys' College, one of the most famous of Scottish schools, draws pupils from professional and business homes. It has elementary and secondary departments and therefore may cover the whole school life of a child educating him from earliest childhood to the time when he leaves for business life or for University studies. A large proportion of the children entering the school at some stage in the elementary department continue to attend at least to 15 years. A number go still farther. The influx of pupils from other schools at the beginning of the secondary school period is considerable but probably does not have the usually noticed effect of increasing the average performance of the boys at the ages concerned. The difference that exists between the children attending this school and those attending ordinary day schools is one of social status. They may <sup>obtain</sup> indicate superior average scores but they are still suitable for the purpose of the experiment in that they present an almost entirely homogeneous group up to 15 years. Beyond 15 years selection operates to complicate matters at the later years.

By/

By the courtesy of the Headmistress of Edinburgh Ladies' College, a secondary school for girls of exactly the same type as George Watson's Boys' College, facilities were given for testing pupils who would correspond in type and age to the boys in the "brother" school.

These three schools provided the children of school age who were tested with the tests, only children from 12 years upwards being included in the investigation.

Adult subjects were obtained from Edinburgh University Psychology Classes, (ordinary and medical) from Edinburgh Provincial College for the Training of Teachers, from Heriot Watt Technical College, from Dunfermline Physical Training College and from the Y.W.C.A. Indian Students' Hostel. Owing to the language difficulty the last named made very unequal scores in the test. The discovery was worth while since it led to the rejection of these results and the results of any foreign subjects tested/

tested in the adult groups \*

### THE TESTS.

The series of tests constructed for application to all subjects was built up on the omnibus-cycle principle. Most of the tests are original though conforming to old principles. Six different types of test were included, namely, analogies, proverbs, mirror letter reading, number series orientation and what may for want of a better name be called 'word penetration'. Each type was represented three times that is to say there were three cycles. At the end of each cycle an odd test was placed. The three odd tests were, logical reasoning, acube test and a comprehensive test. In all 21 tests were given and subsections occurred in many cases/

---

\* I take this opportunity of thanking all who have kindly permitted me to carry out investigations in the schools and colleges and those who have assisted with the testing itself. In particular is my thanks due to Dr Drever for valuable help and encouragement in carrying out the investigation as well as for actually testing the boys of George Watson's College, to Dr Morgan, Principal of E.P.T.C. for many upheavals of the College time-table and to Mr. McLelland, Lecturer in Education at E.P.T.C. for much advice and any statistical knowledge I possess.

cases. Some attempt was made to increase the difficulty of the tests from cycle to cycle but this was done merely on the basis of opinion and, it may be, without precision on this account. Where the type of test permitted a sample test was introduced before the real test question i.e. in analogies and number series. The order of presentation of the tests ~~was the~~ was the same in each cycle. The selected tests were printed on three sides of a double sheet of paper. On the first sheet were the necessary directions, and the time limit for the test. In order to fix a suitable time limit for answering the questions, the test was applied to a number of superior adults and to children of different ages. As the result of their achievements 20 minutes was placed as indicating the time limit.

The Test Sheet is reproduced below. First the directions and then the actual tests are given.

#### PSYCHOLOGICAL EXAMINATION.

TIME LIMIT - 15 Minutes.

NAME/

NAME

DATE OF BIRTH

COLLEGE

DATE

DO not open this Pamphlet until you are told to do so by the Examiner.

On the remaining pages there are a number of problems. In each case you are told exactly what to do. Notice the Instructions carefully.

Do not ask any questions. If you come to a problem that you do not understand, go on to the next problem.

You will be given only 15 minutes. Solve as many problems as you can in the time allowed.

Solve the problems in the order given. Do not skip about on the page.

---

DO NOT TURN THIS PAGE UNTIL YOU ARE TOLD TO BEGIN.

---

1. Each of the lines below, when completed will contain four words, the first two of which are in the same relation to each other as the second two. In each line there is one term missing. Fill in the blank with a word.

EXAMPLES.

(a) King is to Queen as Prince is to  
(The missing word is Princess)

(b) eye is to \_\_\_\_\_ as see is to hear (The missing word is ear).

egg/



egg is to bird	as seed is to
table is to	as window is to glass.
finger is to hand	as is to foot
is to patient	as lawyer is to client
celery is to vegetable	as salmon is to

2. What letter in the word SPECULATION has the same position in the word, counting from the left, as it has in the alphabet?

Answer

- 3.
1. Strike while the iron is hot.
  2. Catch the bear before you sell his skin.
  3. The mouse that has but one hole is soon caught.
  4. It pays to do only one thing at a time.
  5. Desperate people cling to absurd hopes.
  6. Do not attempt the impossible.

Which of the proverbs above means the same as the following: -

- (a) Don't put all your eggs in one basket.

NUMBER

- (b) Make hay while the sun shines.

NUMBER/

## NUMBER

(c) A drowning man will grasp at straws.

## NUMBER

4. If these six letters were seen reflected in a mirror, some would look the same and some would differ. Put a ring round those that would look the same.

H P A L V N

5. Fill in the missing steps in the following number series: -

## EXAMPLES.

(a) 2 4 6 8 10 12 - - (The missing numbers are 14 and 16).

(b) 1 5 2 5 - 5 - 5 (The missing numbers are 3 and 4).

1 3 5 7 9 11 - -

2 2 3 3 4 4 - -

1 4 - - 13 16 19 22

32 30 28 26 24 - - 18

7 6 - 5 7 - 7 3

6. Make a mental picture of some streets with which you are familiar. The streets of the city go North/

North, south, east and west. A man is facing north. He asks for directions to a certain place and is told: "Go straight on till you come to a church. At that point turn east and walk till you come to a school. Then turn south. Your house is on the east side of the street".

Fill in the blank spaces below, using the words right and left instead of the words north, south, east and west, so that the meaning will be the same.

"Go straight on till you come to a church. At that point turn to your            and walk till you come to a school. Then turn to your            Your house is on the            side of the street".

7. Fill in the blanks in the following: -

A is greater than B	C is	than D
A is less than C	therefore B is	than C
B is greater than D	A is	than D

8. Fill in the blanks in the following: -

reward is to hero	as	is to traitor
quarrel is to	as agree	is to friend
poem/		

poem	is to poet	as statue	is to
coal	is to locomotive	as	is to motor car
	is to ignorance	as wealth	is to pover- ty

9. What two letters in the word PREBENDARY have the same number of letters between them in the word as they have between them in the alphabet, and occur in the same order in the word as they do in the alphabet? Write the letters here  
and
10. 1. Sweets grown common lose their dear delight.  
2. Far from the eyes, far from the heart.  
3. A scalded cat fears water.  
4. Every bird likes its own nest best.  
5. There are no gains without pains.  
6. Distance lends enchantment to the view.

Which of the proverbs above means the same as the following: -

- (a) A burnt child dreads the fire. NUMBER
- (b) Familiarity breeds contempt. NUMBER
- (c) Out of sight, out of mind. NUMBER

11. Draw/

11. Draw a line under each of the following words that would look the same in a mirror.

MADAM NOON LEVEL PEEP TOOT MOTTO.

12. Fill in the missing steps in the following number series: -

12 14 13 15 14 16 - -

- - 44 36 28 20 12 4

16 - 15 11 14 10 - 9

1 4 - 16 - 36 49 64

21 18 16 15 - - 9 6

13. When you enter my house you will find a window on your left in the side wall of the passage. When the sun rises it shines straight through this window on to the wall opposite. What direction are you facing when you stand in the doorway and look straight across the street?

ANSWER

14. A three-inch cube, painted red, is sawed into inch cubes. How many of the inch cubes have paint on two faces only?

ANSWER

15. Fill in the blanks in the following: -  
caress/



caress is to as blow is to anger  
 vestibule is to house as is to book.  
 not is to ton as but is to  
 before is to behind as future is to  
 aspect is to space as phase is to

16. What are the two letters in the word EQUILIBRIUM which have the same number of letters between them in the word as they have between them in the alphabet, but whose order in the word is the reverse of that in the alphabet?

ANSWER and

17. 1. A stitch in time saves nine.  
 2. Untempted virtue is easily retained.  
 3. Birds of a feather flock together.  
 4. A small leak will sink a ship  
 5. Let the cobbler stick to his last.  
 6. Persistence knows no defeat.

Which of the proverbs above means the same as the following: -

- (a) Nightingales sing their own songs best.

NUMBER.

- (b) Reputation may be ruined by a word.

NUMBER/

## NUMBER

(c) Constant dropping wears away stones.

## NUMBER

18. A clock seen in a mirror appears to show the time twenty-five minutes past eight. What is the real time? ANSWER

19. Fill in the missing steps in the following number series.

24 27 - - 32 35 36 39

2 - - 8 12 17 23 30

39 34 30 - - 16 12 7

4 8 10 20 22 - - 92

81 27 54 - 36 12 24 -

20. If a boy stands on his head with his face to the south, where will his right hand point?

## ANSWER

21. Read: "The poetry of Milton differs from that of Dante in that the images which the latter uses, being employed, not for ornament, but simply to make the meaning clear, speak for themselves, whereas those of the former have a value which depends less on what they directly represent than on what they remotely suggest" /

suggest".

ANSWER - 1. Which of the two poets mentioned  
employs images that speak for  
themselves?

2. For what purpose does Dante em-  
ploy images in his poetry?

3. On what does the value of Milton's  
images depend?

---

No further directions than those indicated  
were given. The time was regulated by the use of a  
stop watch

#### MARKING.

---

In this initial stage of the investigation  
the marks were more or less arbitrarily fixed to the  
various questions, but, since it was a time limit  
test, with a certain regard for the time taken to solve  
the problems. For each single analogy completed  
one mark was given. The questions with no subdivi-  
sions viz. Nos 2, 4, 9, 11, 13, 14, 16 and 18 - re-  
ceived four marks each. The question on logical  
reasoning/

reasoning also received four marks in spite of subdivision within it. Partial credit was allowed, two marks being given if two spaces of the total three were correctly filled. No marks were allowed if only one space was filled. In the proverbs tests each of the three subdivisions valued two marks making a total<sup>of six</sup> for each of the three questions, 3, 10 and 17. In the number series completion tests one mark was given for each series completed, that is 5 marks were obtainable for each of questions 5, 12 and 19. Question 6 (orientation) was most easily marked by giving partial credit for each direction noted and the same applied to question 21. These marks happily produced a total of 100, a convenient number to work with.

#### KEY TO THE TEST - CORRECT SOLUTIONS

---

As far as possible verbal tests were chosen so that only one response would be suitable. In certain cases it was necessary to give credit to the ingenuity of children but very seldom was this done and only in cases where it was impossible to discredit a naive response. The number series questions have one solution only ( but there is one answer to all in spite of some teachers' disbeliefs.)

1. plant: wood: toe: doctor: fish:
2. i.
3. 3, 1, 5.
4. H: A; V;
5. 13, 15: 5, 5: 7, 10: 22, 20: 7, 4.
6. Right; right; left.
7. Greater; less; greater.
8. Punish or punishment: enemy or foe: sculptor: petrol: knowledge.
9. P and Y
10. 3; 1; 2
11. TOOT
12. 15, 17; 60, 52; 12, 13; 9, 25; 12, 10
13. North
14. 12
15. Love; preface or introduction: tub: past: time.
16. M and U
17. 5; 4; 6;
18. Twenty five minutes to four.
19. East
20. Dante: to make the meaning clear: on what they remotely suggest.

Age/



## AGE GROUPING.

---

The principle adopted throughout in regard to grouping children according to age is to classify from half year to half year. For example, 13 years is written always to express the group including children from  $12\frac{1}{2}$  to  $13\frac{1}{2}$  years and so with the other age groups.

## PROCEDURE.

---

The testing was carried out during the months of March, April and May 1924. The first test was only a tentative approach to the subject. All the subjects in the James Clark School in Edinburgh Ladies' College, Edinburgh University Psychology Class and in Edinburgh Provincial Training College were tested by the investigator. The subjects in Dunfermline Physical Training College were tested by Miss Young, assistant in Edinburgh University Education Department, and the subjects in the Heriot Watt Technical College by Mr Muir, Lecturer in Engineering and fellow research student in the department of Psychology. Owing the fact that George Watson's/

Watson's Boys' College drew boys and Edinburgh Ladies' College girls from the same family it was considered wise to test both schools on the same day. Accordingly Dr Drever tested the boys and the investigator tested the girls.. The same procedure was strictly observed in all testing and all who tested had considerable knowledge of testing difficulties. All test results were handed over to be dealt with by the investigator personally.

In the James Clark School the children were tested in their classes, each class singly. The testing occupied two entire school days. Children who were tested during the first day were put on their honour to have no conversation whatever about the tests. The necessity for this was probably not great because the tests were not for the most part such as could be readily remembered. Nevertheless the only precaution which could be taken, was taken.

In Edinburgh Ladies' College the same method was followed and also in the adult groups tested. Only in G.W.B. College was the work facilitated by the grouping together of the boys in two very large masses in the school hall. Little over an hour sufficed/

sufficed for the testing of the whole number. Whether the size of the groups tested makes any difference to the results has not been ascertained. It has been suggested by some investigators that individuals should not be tested in groups that are too large. There is no outward evidence that the G.W.B.C. results suffered on account of the size of the groups. On the contrary, of all the groups tested, it presents the most satisfactory and regular results, so that it might be asked if the grouping of the pupils in a mass, regardless of class differences, does not act as a stimulus to good achievement. It is not desirable from the point of view of testing that this view should hold good, because the wish is not to create an atmosphere that is unusual. But another point is noteworthy. Any subjective influence arising from the personal equation of the tester, which is well-known to be variable, would be less likely to influence unequally the results of one or two large groups than the results of, say twenty such.

# JAMES CLARK SCHOOL. (Technical)

The test was applied to 446 children of both sexes, from 12 to 16 years of age, inclusive. Of these 239 were boys and 207 were girls.

The Table below shows the numbers of each sex tested at the different ages;

12 years =  $11\frac{1}{2}$  Years =  $12\frac{1}{2}$  years and so on.

Age	12	13	14	15	16	Total
Boys	12	84	100	35	8	239
Girls	21	60	90	30	6	207
Total at Each Age	33	144	190	65	14	446

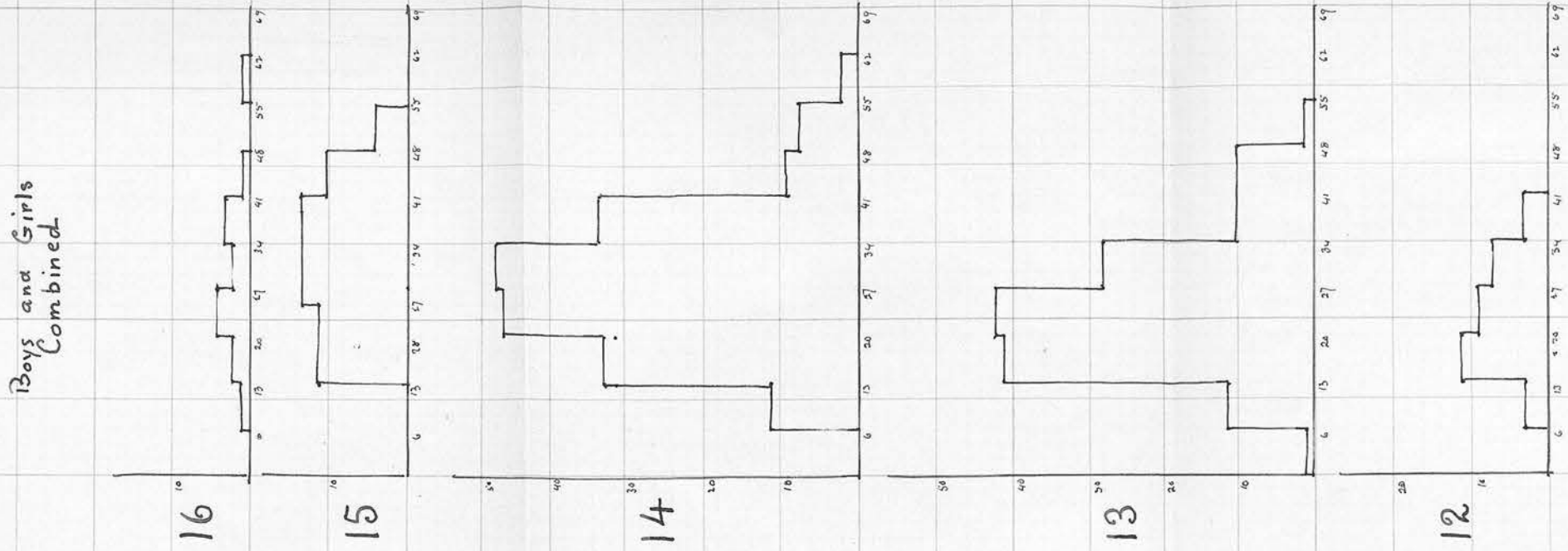
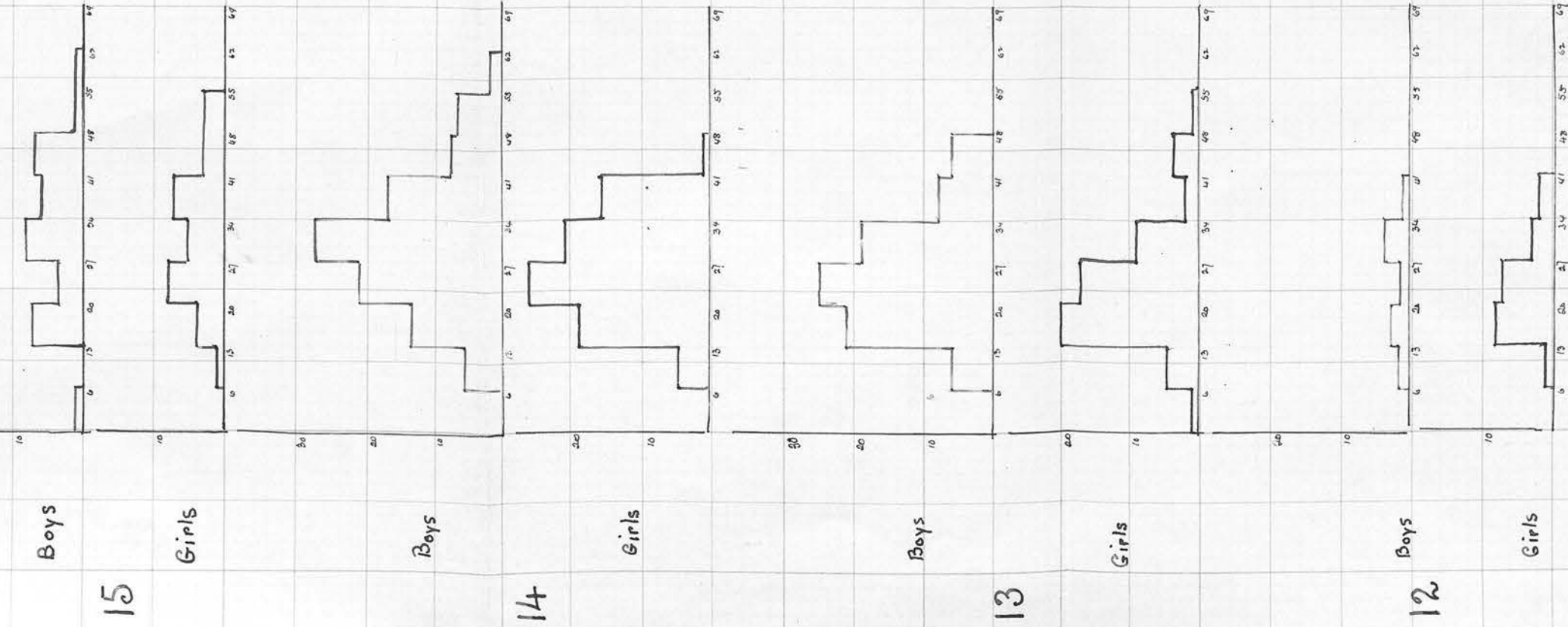
The first thing that meets the eye is the inequality of numbers at each age. A particular paucity of subjects is evident at 12 years and at 16 years. This can easily be explained. Children coming at first to the school are more likely to fall into the 13 year old group because it extends from  $12\frac{1}{2}$  to  $13\frac{1}{2}$  years. Those in the 12 year old group are most probably of the 'young bright' variety. Again, the normal course at the school is of three years' duration so that subjects in the 16 year old group are probably duller or otherwise retarded pupils/

pupils. This explains the small number at the uppermost age group. There is one further fact to be noted. It appears that, although the normal course of study extends over a period of three years, a considerable number of children leave at 14 years of age. Their purpose in the school presumably is only to obtain the benefit of specialised education during the last two years of compulsory school attendance. The school has to tolerate the presence of such children although its concern is not primarily their education. The table above reveals the fact of their withdrawal very clearly. The matter is interesting from the point of view of this experiment, because selective factors might obscure the actual course of average mental development from 14 to 15 years. Moreover this may be supported as the result of subsequent revelations. This same school was retested after an interval of nine months and if it can be assumed that the second test was as good a test of intelligence as the first, it appears that the selection on the basis of ability is the chief factor. Of those children who were tested at 14 years in March, the majority of whom had presumably entered the 15 year old age group in December, as/



# I<sup>st</sup> TEST

## JAMES CLARK SCHOOL



Scale 3 units horiz. = 1 class interval. (7 units)  
 5 units vert. = 10 subjects

as many dull as bright had returned. The age grouping from half year to half year coupled with the interval of less than a year obscures this to some degree, but a penetrative analysis reveals that the dullards are eliminated more than the able at this point. The selection by 'ability' goes hand in hand with 'opportunity' and other external factors. Another difficulty in judging the results at 15 years is due to the small numbers available for testing.

The second thing that the Table reveals is that the number of boys and girls at each age group is approximately equal, a circumstance which renders the comparison of the sexes easier.

#### J.C. DISTRIBUTIONS.

---

The distribution of scores for each age was drawn up to facilitate calculations and give a pictorial representation of the results. These distributions are shown in miniature form on the opposite page. They are, on the whole, fairly regular, except at the points where skewness might be expected. The overlap from age to age is considerable/

# BOYS

Age	Mean	St.Dev <sup>n</sup> .	Median	P.E.of Mean	No.
12	23.25	9.3	22	1.9	12
13	25.66	9.1	25.3	.66	84
14	29.95	11.2	30.55	.75	100
15	33.74	12.6	34.8	1.4	35
16	32.83	13.7	28.5	.	8

# GIRLS

Age	Mean	St.Dev <sup>n</sup> .	Median	P.E.of Mean	No.
12	23	7.2	23.16	1.05	21
13	23.71	9.8	22.9	.86	60
14	26.24	8.19	26.56	.58	90
15	30.13	10.2	28.5	1.2	30
16	26.16	10.2	25.5	.	6

# ALL SUBJECTS.

Age	Mean	St.Dev <sup>n</sup> .	Median	P.E.of Mean	No.
12	23.12	8.26	24.5	.97	33
13	24.8	9.45	24.5	.53	144
14	28.7	10.22	28.4	.49	190
15	32.1	11.9	32.7	.99	65
16	30.4	12.6	27.25	.	14

considerable. The girls' distributions are less regular, if anything, than the boys'. The distributions in themselves point, though not too clearly, to a slight increase in average growth throughout the school period. They require to be interpreted in the light of known facts regarding the subjects.

The chief interest of the results being in the average performance of the subjects at each age and their general variability, the means and standard deviations were calculated for each age, taking boys and girls separately, and together. For subsidiary purposes the median and Probable Error of the Mean were also derived.

The results are presented in the table below.

#### ANALYSIS OF RESULTS.

It is assumed in the type of investigation that age 14 represents what in the course of a year age 13 will become. Without the assumption the comparison of successive age groups means little. Comment has already been made on the limitations of the/



207

the group method of testing so that further remarks on this matter are unnecessary.

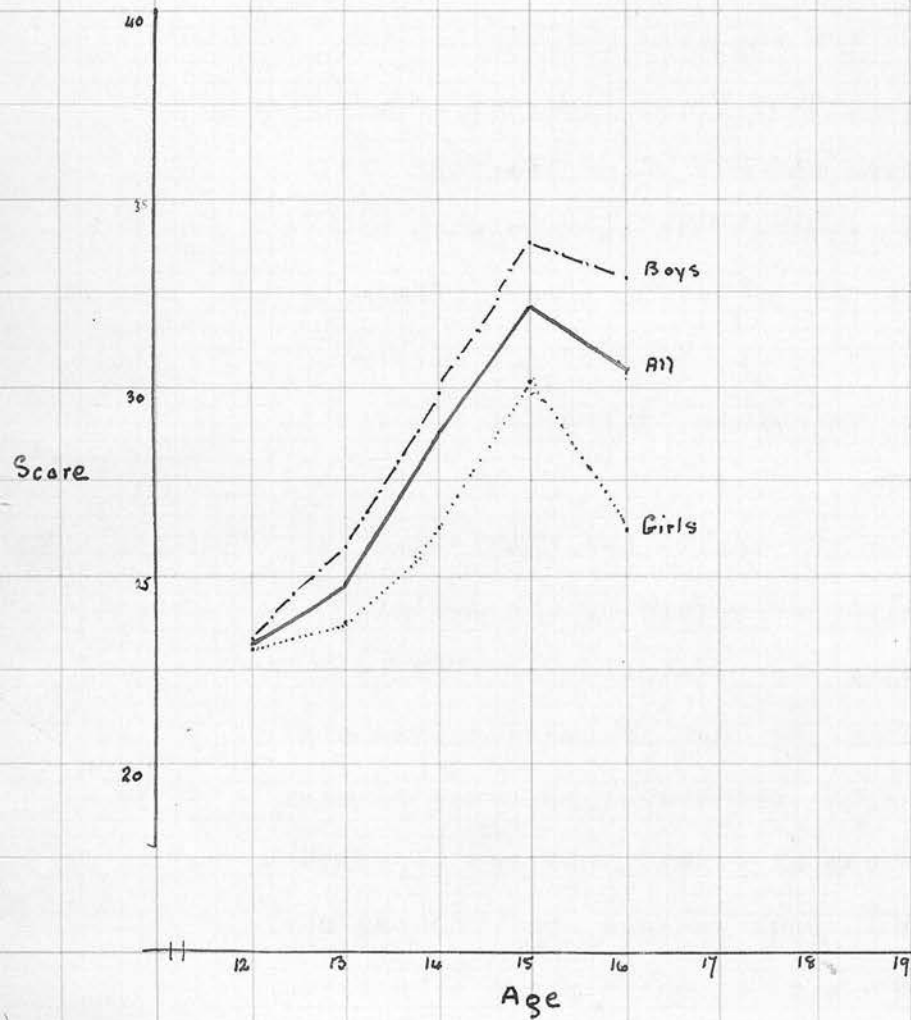
The tables show a steady increase in the average performance of both boys and girls during the ages 13, 14 and 15. There is at all points a slight superiority of boys noticeable. The results at 12 and 16 years serve only to bring out what was previously noted regarding the composition of the groups at these ages, namely that the 12 year olds are few and bright while the 16 year olds are few and dull.

Parallel with the increase in average score from age to age is an increase in the variability of each group, more consistently apparent in the boys' results. The boys show a tendency to be more variable than the girls at all ages except one. The peculiar instance is at 13 years, where the girls' results show a variability of 9.8 against the boys' 9.1 at the same age. The period at which the difference occurs is significant at least. The two facts increase in variability from year to year and greater variability among boys are in keeping with general belief that variability increases in going towards the adolescent period and that boys are more variable/



# I<sup>st</sup> Test

James Clark School



Scale { 5 units horiz. = 1 year  
2 units vert. = 1 mark

variable on the whole than girls.

From the tables it is a simple matter to construct the curves of average growth from age to age. Accordingly these curves have been drawn and are placed opposite. The red line represents the combined results, the dotted <sup>green</sup> black line the girls' results and the 'stroke and dot' line the boys' results. These curves taken in conjunction with the table of results make the situation clear. There is considerable similarity in the form of the curves of both sexes, though the girls' curve falls a little lower than the boys'.

Neither from the tables nor the curves is it easy to draw conclusions regarding the age of mental arrest. Viewing the results superficially one might be tempted to say that 15 years represents the maximum point reached and therefore seems to mark the limit of average mental growth. At age 16 there is no advance on the 15 year average, but instead a slight retrogression - a point which might give still further support to the view.

The data require closer analysis.

#### BOYS.

To begin with the boys' results, it is seen from the tables that there is an average at 12 years of/  
 of/

of age which is higher than would be expected from average subjects of that age, <sup>tested</sup> the children<sub>1</sub> were undoubtedly bright. Accordingly the interval between 12 years and 13 years is somewhat smaller than it should be; or, looking at it from the inspection of curves, the ascent of the curve line is steeper than it ought to be. The intervals between 13 and 14 years and between 14 and 15 years seem to be marked by regular average growth. No selective factors enter in the first group so that it is probable that the growth is real. The reliability of the means suggests that they are trustworthy. Between 14 and 15 years interpretation is not so simple. The average growth indicated here barely equals that in the previous interval. Moreover, at 15 years selection begins to operate freely, so that the average score at that age is not indicative of normal growth alone but is due also to the fact that superior ability has contributed to it. These two facts, taken together, point to a diminution in average mental growth in the interval between 14 and 15 years. The 'barely equal growth' would not itself suffice to suggest a decrease but coupled with selective influence it makes no other interpretation possible.

At/

At 16 years there is a positive decrease which in itself is not remarkable since the 16 year olds are probably dull. What is interesting is that even the decrease surpasses the 14 year old level so that the few dull 16 year olds, but slightly inferior to the 15 year old level, are markedly superior to the average 14 year olds.

#### GIRLS.

The analysis of the girls' results follows similar lines. The 12 year old girls barely reach the average of the 13 year olds. The boys aged 12 years surpassed the following age. The interval between 13 and 14 years indicates a certain amount of average growth, which is, however, less than indicated in the boys' results. Between 14 and 15 years the growth which the actual figures and curve reveals is greater than in the previous interval. Since there is every reason to believe that the averages at 13 and 14 years are reliable, some reason for the increase must be found. It is without doubt due to selective factors entering at 15 years and tending to obscure the true course of growth. The selection that operates may easily be more rigid in the case of the girls than in the case of the boys/

boys , but it cannot be proved. At 16 years the girls represented attain only the height of the 14 year olds but the numbers are insufficient for any reliability to be attached to the figures.

#### ALL SUBJECTS.

The results obtained by combining the scores of boys and girls show the same characteristics as the separate results. The interval between 14 and 15 years misrepresents the true average growth between these years. There is in reality every indication of a decrease in average growth after 14 years. Where this decrease becomes a minimum cannot be said but without doubt it is not later than 16 years. It may be even sooner.

A comparison of average Scores in two groups does not tell much regarding the differences in groups as a whole. If it is assumed that the distributions are approximately equal the measure of variability may suffice to overcome the difficulty. But, for adequate and satisfactory comparison it is safest to indicate the percentage of the one group reaching or exceeding the median of the other. In this case it has been chosen to represent the/  
the/



the percentage of each group reaching or surpassing the median of the median below.

The tables below give the actual percentage.

	BOYS	GIRLS	ALL
Percentage of 13 year olds passing median of 12 year olds	41.6	57	54.8
Percentage of 14 year olds passing median of 13 year olds	50	62	57.8
Percentage of 15 year olds passing median of 14 year olds	60	62.8	60
Percentage of 16 year olds passing median of 15 year olds	"	37.5	37.7

It is to be feared that the table is no more illuminating than the other results. The difficulty is that there are too few age groups represented and even such as are included require to be interpreted through other known facts beyond the scores themselves. The boys' table suggests a superiority of 15 year olds over 14 year olds which is still significant but the table extends no further. At the age below, no superiority exists so that the suggestion is that the superiority at the year above is due to selection. But the girls' table does not support such a finding and moreover, conclusion/

conclusions relating to the 15 year old group must always be made in remembrance of the small number tested there. Up to 15 years the girls' distribution or group is progressing. Probably the figure 62.8 is due partly to selective influences, but some part of it goes to indicate a general superiority of the 15 year old group over the 14 year old group. The percentages bearing on all the results taken together are once more obscured by the selective agencies at work from 15 onwards. As suggested this method of expressing the difference between the successive groups serves merely, in this case, to support the findings previously noted, namely that up to 14 years at least growth is progressing. After 14 years there is some indication of a slight average decrease in growth and after 15 years this decrease undoubtedly becomes marked, though the inadequacy of the data at hand and the limited range of ages tested render the task of fixating the limit with accuracy a very difficult one. Still, weighing the increase in average growth against the selective force whose influence is generally limited to the exclusion of the lowest intelligence it appears that no concession can be made, therefore to DOLL'S proposal that the upper limit of the growth of intelligence is at  $13\frac{1}{2}$  years. Beyond this nothing further need be said here.

### ALTERATION OF THE TIME LIMIT.

It was incidentally discovered while testing the JAMES CLARK School that an adult who had considerable practice with mental tests was able to complete the test sheet and obtain nearly the maximum score in slightly less than 20 minutes. The circumstance was unfortunate, but it was deemed advisable to alter the time limit to 15 minutes, since many adults were to be tested later. Therefore, in subsequent applications of the test only 15 minutes was allowed.

-  $14\frac{1}{2}$  age group. (This is a consistent happening as the second test shows the same increase). The time seems to be a little late for an influx of pupils, falling as it does between 13 and 14 years. But there it is. It is probable that the new arrivals do not materially alter the general average scores for the group. Later comparison of children of different social status will make the reason clear,

At 12 years the children tested are again of the young bright variety, being advanced for their years since 12 they have a place in the secondary school. But from 13 years onwards to beyond 15 years it is safe to assume that selective factors do not operate. What distinguishes these children from other children, is the difference in social status. After 15 or 16 years the selection of able pupils occurs and continues to operate till the end of the school period. The paucity of subjects at the upper end of the table is the result of the rigorous selection.

In the body of the table the numbers at each life age are sufficient to yield satisfactory results from the application of the test. Further in the three age groups 14 15 and 16, the number of subjects is approximately equal. The distributions of scores/

1<sup>st</sup> test

George Watson's Boys College.

No  
of  
Subjects

18

17

16

15

14

13

12

Score

Scale.  $\begin{cases} 2 \text{ units horiz.} = 1 \text{ class interval (7 points)} \\ 5 \text{ units vert.} = 20 \text{ subjects} \end{cases}$



scores are placed opposite. They are for the most part regular, excepting in places where regularity cannot be expected on account of the small number of subjects tested. The overlapping of the successive age groups is so great as to render analysis impossible without a knowledge of other factors. The distributions reveal a slight tendency to increase up to 16 years and thereafter little improvement can be observed. The fact may be due only to the smaller numbers tested after 16 years, but it is noteworthy at least that after 15 years when selection on the basis of ability first appears there is no corresponding tendency for the scores either to increase noticeably at the top end of the distribution or to shorten the 'tail' of the curves at the other end. The schools of this type, as intimate knowledge shows harbour many who would profit more by being in entirely different spheres. Opportunity alone ensures their presence always at the later years.

The table below shows the results obtained for each age group tested:-

Table/

AGE	MEAN	ST.DEV. <sup>n</sup>	MEDIAN	P.E.OF MEAN	NOS.
12	37.6	(12.8)	38.3	1.21	20
13	33.42	9.8	33.6	.76	74
14	34.72	11.76	34.1	.74	113
15	33.46	11.97	33.8	.78	105
16	37.7	11.9	38.1	.76	109
17	40.87	11.9	40.7	.95	80
18	46	11.37	46	1.21	40
					6

#### ANALYSIS OF RESULTS.

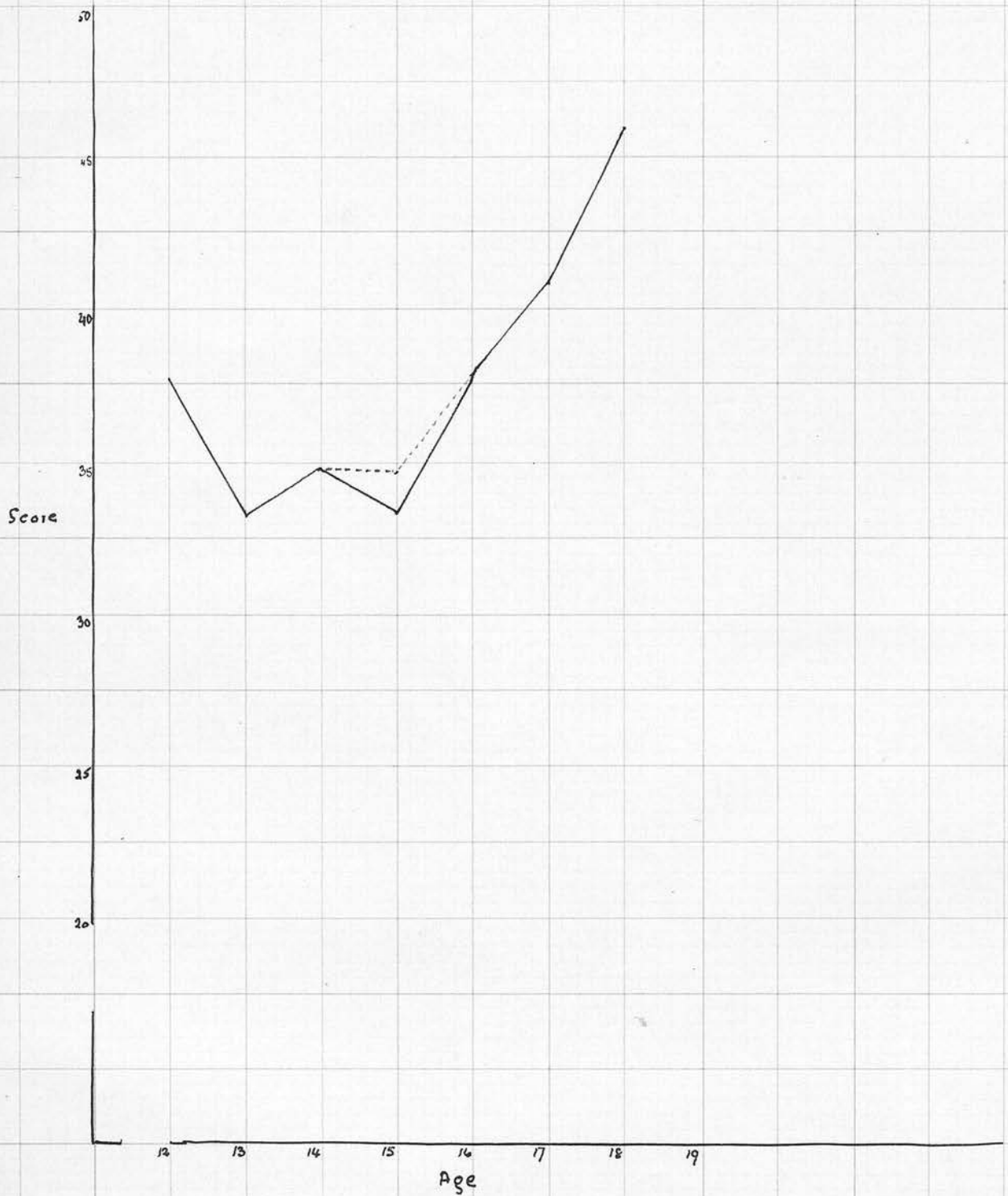
---

The figures are not at first sight reassuring. The large average at 12 years is explained by the selection of subjects and the accompanying variability of the group is not reliable on this account. The average score at 13 years is coupled with a variability measure which suggests a homogeneous group. The distribution confirms this suggestion. From 14 years onwards there is fairly steady increase in average score, with a 'reversal' at 15 years which looks ominous. At the same time the groups are remarkably similar in variability. At the later ages this/

this may be accidental but it is not without interest, and comparison of one group with another is rendered simpler on this account. The interval between 13 and 14 years is marked by a very slight increase in average score, which, if it be taken to indicate average growth, reveals that progress is very slight. At 15 years, however, the score decreases, a fact which might suggest that 15 year old subjects are on the average only as good as 14 year old subjects. To study averages alone is insufficient. The marked difference in the variability of the two groups 13 and 15 years suggests that a further analysis of the group scores is necessary. The disturbing factor at 15 years immediately becomes apparent. A glance at the distribution shows that extreme measures at both ends influence the average. The group contains one very 'intelligent' boy and five very 'stupid' boys. The tail of the curve is unduly lengthened almost to reach zero ability. The elimination of these five cases heightens the average to 34.66, that is to say, the 15 year olds then show an average ability which approximates to the average ability of the 14 year olds. After 15 years selection enters to eliminate many/

1<sup>st</sup> Test.

George Walson's Boys' College



Scale. { 5 units horiz. = 1 year  
2 units vert. = 1 mark

many of the children of poorer mental ability. The average score at 16 years is large enough to show that the increment from 15 to 16 years is due mainly to this factor. The fact that there is no appreciable difference between the 14 year group and the 15 year group lends support to the opinion that it is all due to selection. The same average increment characterises the next interval, while at 18 years the average score is much larger than at the previous age.

The growth curve opposite, which is derived from the average scores, exhibits the results in clearer form. The most noteworthy feature is its flat appearance from 13 to 15 years. The dotted line in violet shows what the course of the curve would be with the elimination of the extreme scores at 15 years.

The addition of the table showing the percentage at each year reaching or surpassing the median of the previous year, does not in this <sup>case</sup> year add much to what has already been discovered. Its interpretation involves a reference beyond the figures themselves.

Table/



	Percentage of 13 year olds reaching or surpassing the median of the 12 yr. olds	27
" " 14 "	" " " " " " " " " " " "	46.9
" " 15 "	" " " " " " " " " " " "	49.5
" " 16 "	" " " " " " " " " " " "	61.4
" " 17 "	" " " " " " " " " " " "	57.5
" " 18 "	" " " " " " " " " " " "	65

At 12 years the figure is not unexpectedly small but at 13 and 14 years each group barely manages to equal the preceding group. When the difference in variability between the 13 year and 14 year group is remembered the low figure is not surprising. The figure 49.5 supports the conclusion arrived at previously that as far as can be determined from the data there is no appreciable increase at 15 years over 14 years and that, further, from 13 years onwards the average growth is so small, as to indicate that the upper limit of the growth of intelligence is revealed soon after this point. The tables, the curves and the percentage figure indicate 14 years it is true, but it is not easy to accept the results for the 15 year old group as being satisfactory. The mean has if anything a greater unreliability than it ought. But the chief reason for dubiety is that the results from Edinburgh Ladies' College, which are expected to follow the same average course, point to a slight continuance of growth even beyond 14 years. The James Clark results, as far as they permitted analysis seemed to point to a similar conclusion. Care must be exercised in selecting any fixed level as marking the limit./

limit. The distributions reveal the fact that the overlapping from age to age at all periods within the scope of this investigation is so great as to swamp the differences from age to age. The selected 12 years old boys tested, have as high an average score as the average 16 years old boys. At 12 brightness explains all, at 16, however brightness and dullness intermingle. The fact is certainly worthy of note.

#### EDINBURGH LADIES' COLLEGE.

---

The girls at Edinburgh Ladies' College were tested in their classrooms, and not in large groups. A certain number of the older pupils were not tested. They were specialising in science and were out on a pre-arranged botanical excursion. In all 428 pupils were tested during part of one school day. The experimenter had to adopt the plan of starting one class, noting the exact time when it must stop and proceeding to another class. In this way a few classes were kept going at once. Familiarity with the school made the matter simple. All the classes were tested before the lunch interval so that the danger of communication/

communication was avoided.

The numbers tested at each age are given below:-

AGE	12	13	14	15	16	17	18	19	20	
IN°	34	72	85	84	69	49	24	8	3	428

At 12 years and above 17 years the number of girls tested is too small to permit of adequate conclusions being drawn regarding these groups. In the middle section the numbers are larger though only at 13 years as large as the parallel numbers in the boys school. There is no marked influx of pupils from  $13\frac{1}{2}$  to  $14\frac{1}{2}$  as was shown in the  $13\frac{1}{2}$  to  $14\frac{1}{2}$  age groups of G.W.B.C It is a noteworthy feature. It happens that the influx of girls from other schools occurs considerably earlier, and is not marked at any stage. The junior school collects the pupils who continue to attend at least to the 'intermediate' Stage. Only a sprinkling of new pupils enter at the  $12\frac{1}{2}$  to  $13\frac{1}{2}$  or  $13\frac{1}{2}$  to  $14\frac{1}{2}$  age groups. After 16 years the selection of superior intelligences begins and increases till the end of the school period when probably/

probably mostly bright children remain.

In the table below the results obtained from each age group are collected:-

AGE	MEAN	ST.DEV. <sup>n</sup>	MEDIAN	P.E.OF MEAN	N°
12	30.76	8.4	29.78	.96	34
13	31.08	9.8	31.1	.79	72
14	32.6	10.5	32.74	.87	85
15	35.83	11.9	35	.90	84
16	38	9.45	37.1	.81	69
17	40.1	11.9	39.5	1.1	49
18	44.89	11.3	44.33	1.23	24

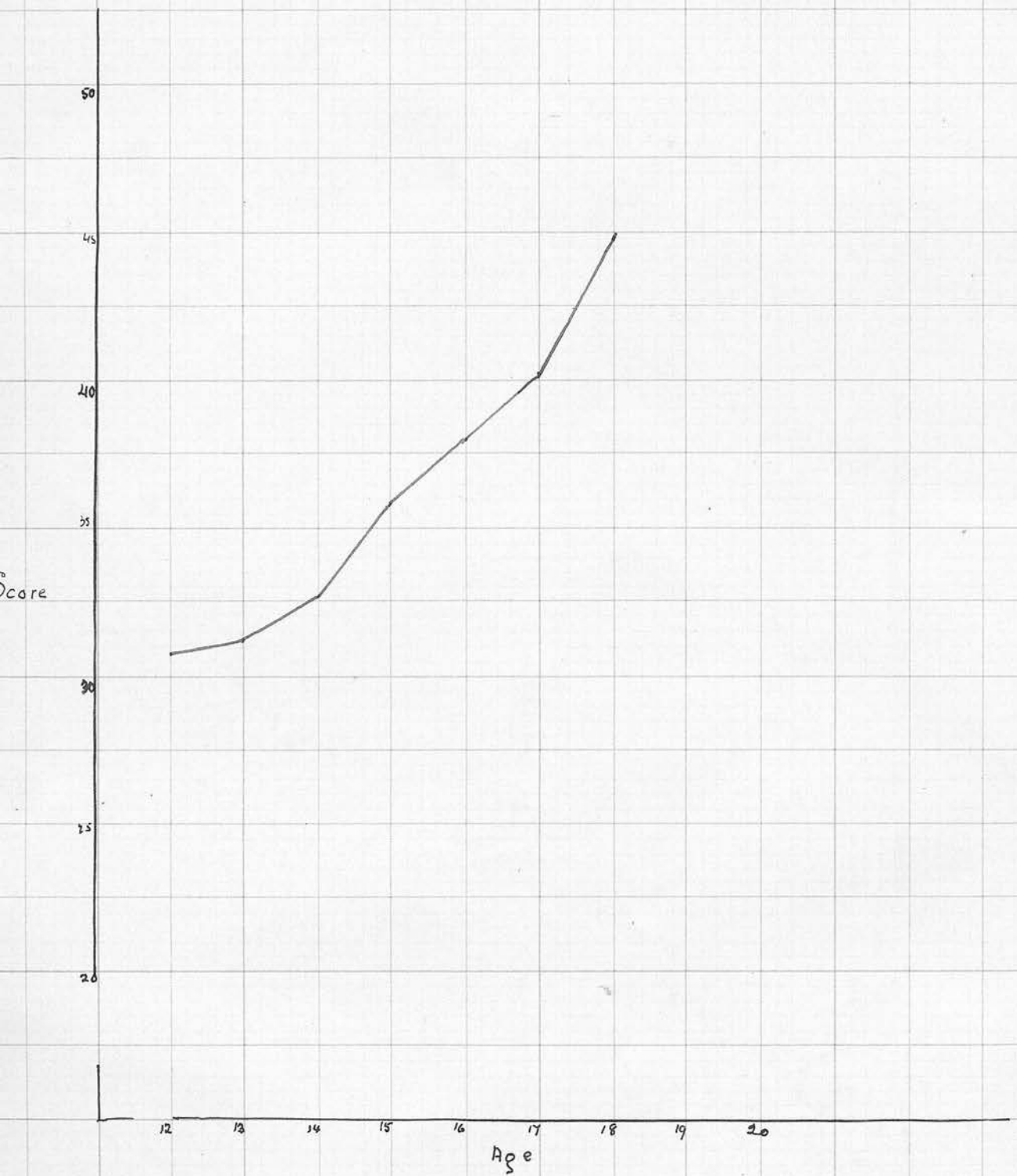
#### ANALYSIS OF RESULTS.

At 12 years, as before, selection tends to heighten the average. The bright children of this age were tested but also a few from the pre-secondary school stage. The result of the combination is that the average score which the 12 years old girls make is not so high as that which was indicated in the corresponding group of boys. Great reliability is not attached to the figures on account of the small number tested. The 13 years old group shows a slight average superiority over the 12 years/



years old group which would probably have been more marked if selection had not entered at the lower year. From 13 to 14 years the average increment is small, from 14 to 15 years it assumes a larger size. Some explanation of the increase in average growth is necessary. At this point the main concern is to explain the increase relative to the whole age series. Two possible explanations present themselves. Either there is evidence of selective factors at work, or the apparently large average increase is due to usually small average scores at the lower years. Neither explanation is improbable. Though, in dealing with the George Watson's Boys College results it was assumed that the selective factors did not enter at 15 but at 16 years, as the figures are used here, there are good reasons for believing that selection on the basis of ability begins somewhat earlier with the girls than the boys. A comparison of the first 14 year old group with the second year 15 year old group revealed that those who returned to school were, if anything slightly superior in average ability. But the boy's group showed no such happening. It is conceivable that the girls of lower ability do tend to be eliminated. The elimination is not marked, only/

I<sup>st</sup> Test  
Edinburgh Ladies' College



Scale. { 5 units horiz. = 1 year  
2 units vert. = 1 mark

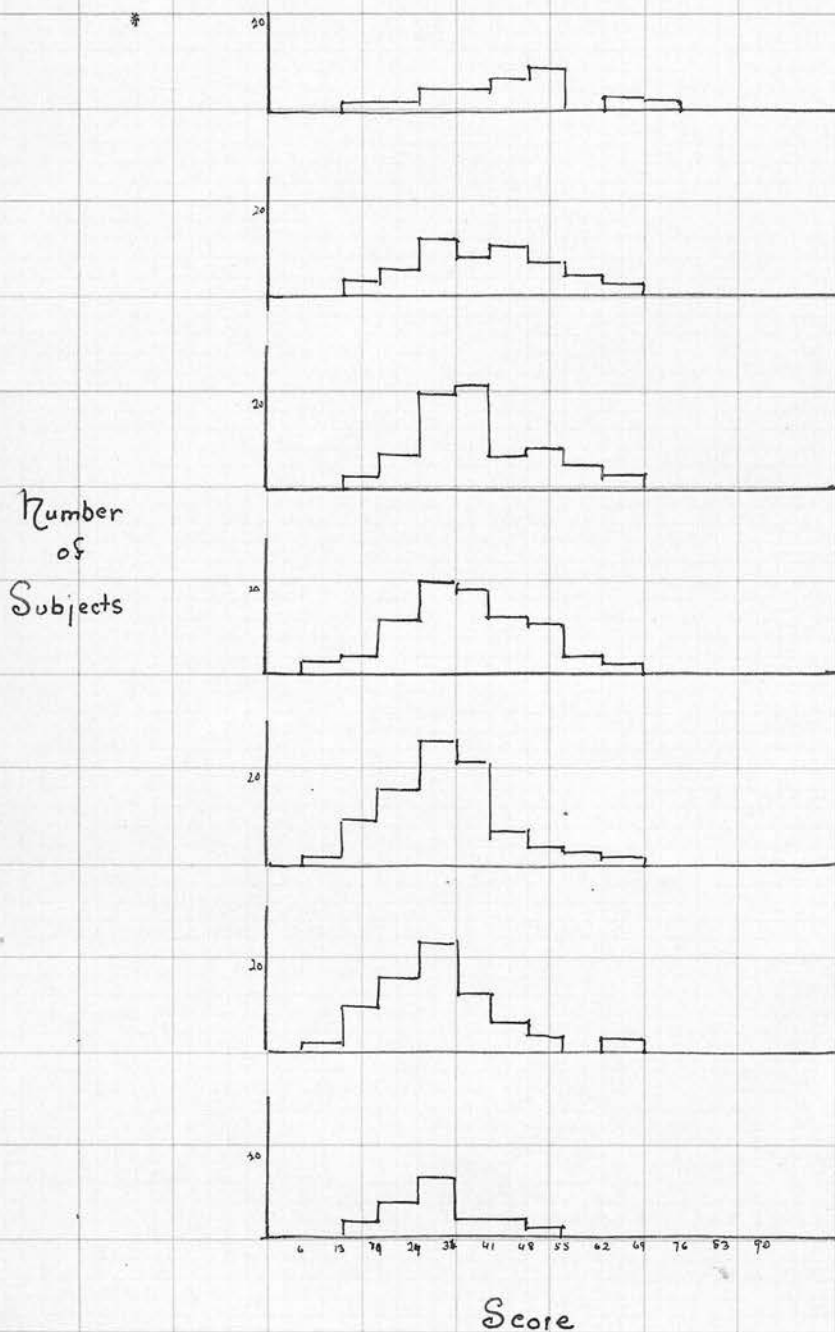
only observable.

Perhaps a better explanation offers itself when it is remembered that probably the girls of 13 and 14 years might naturally be expected to exhibit lower average scores than the boys of the same age. Unfortunately, the ages do not extend low enough to show if this suggested explanation is the real explanation, but confirmatory evidence lies in the boys' results. The average score at 15 years has already been commented on and dissatisfaction expressed regarding it. It is not surprising now that it should occur at that point and it is naturally later than with the girls. The puberty question offers a satisfactory explanation of the anomalies at different periods in each sex. But interpretation must also bear in mind the unequal number of boys and girls represented at each age.

Beyond 16 years selection is rigorous. The increase in average scores may be entirely a misrepresentation of normal average growth. With regard to the variability of the groups, it is noteworthy at least that after 14 years all age groups show approximately the same standard deviation. The boys' results revealed the same characteristic. The small variability at 16 years, which the distribution of scores/

# I<sup>st</sup> Test

## Edinburgh Ladies' College.



Scale.  $\begin{cases} 2 \text{ units horiz.} = 1 \text{ class interval (7 units)} \\ 5 \text{ units vert.} = 20 \text{ subjects} \end{cases}$

scores makes clear may be due to the elimination after the so called 'intermediate' stage.

The growth curve based on the results and drawn opposite presents a fairly regular appearance. The trend of the curve after 14 years is deceptive and not indicative of average but of 'selected' growth. If the most reliable scores are ~~found~~ at 13, 14 (and 15?) years, then there is no danger of misinterpretation. The growth curve shows a tendency to become parallel in the earliest stages but the tendency is counteracted by the unduly large increments at the later ages.

In the table below the percentage of each group reaching or surpassing the median of the group below indicates the general change from one year to another in a more concise form than the distribution (opposite) from which the results previously analysed are derived.

Table/



Percentage of 13 year olds reaching or exceeding the median of the 12 year olds.												48.8
"	"	14	"	"	"	"	"	"	"	13	"	50.6
"	"	"	15	"	"	"	"	"	"	14	"	55.9
"	"	"	16	"	"	"	"	"	"	15	"	44
"	"	"	17	"	"	"	"	"	"	16	"	57.1
"	"	"	18	"	"	"	"	"	"	17	"	62.5

From the table there is no possibility of fixing the average upper limit of intelligence beyond 15 years. The percentage of 16 years old reaching or exceeding the median of the 15 years old is only 44%; that is to say the measure of change is all in favour of the lower group. At 16 years, moreover, selective forces are operating so that instead of a decrease an increase might be expected. The boys' results show such an increase when selection enters. From 13 years to 14 years the change is almost negligible and from 14 years to 15 years the increase is marked. Taking into consideration all the known facts concerning the condition of the subjects, and the selective factors in operation it seems safe to conclude that the figure 55.9 at 15 years is not due to selective factors so much as to a 'revival' of ability in the girls generally. Selection may contribute its share. The decrease at 16 years is an indication that average progress is non-existent. The heightened percentage at 17 and 18 years can be explained as being solely due to selection of able pupils at these ages.

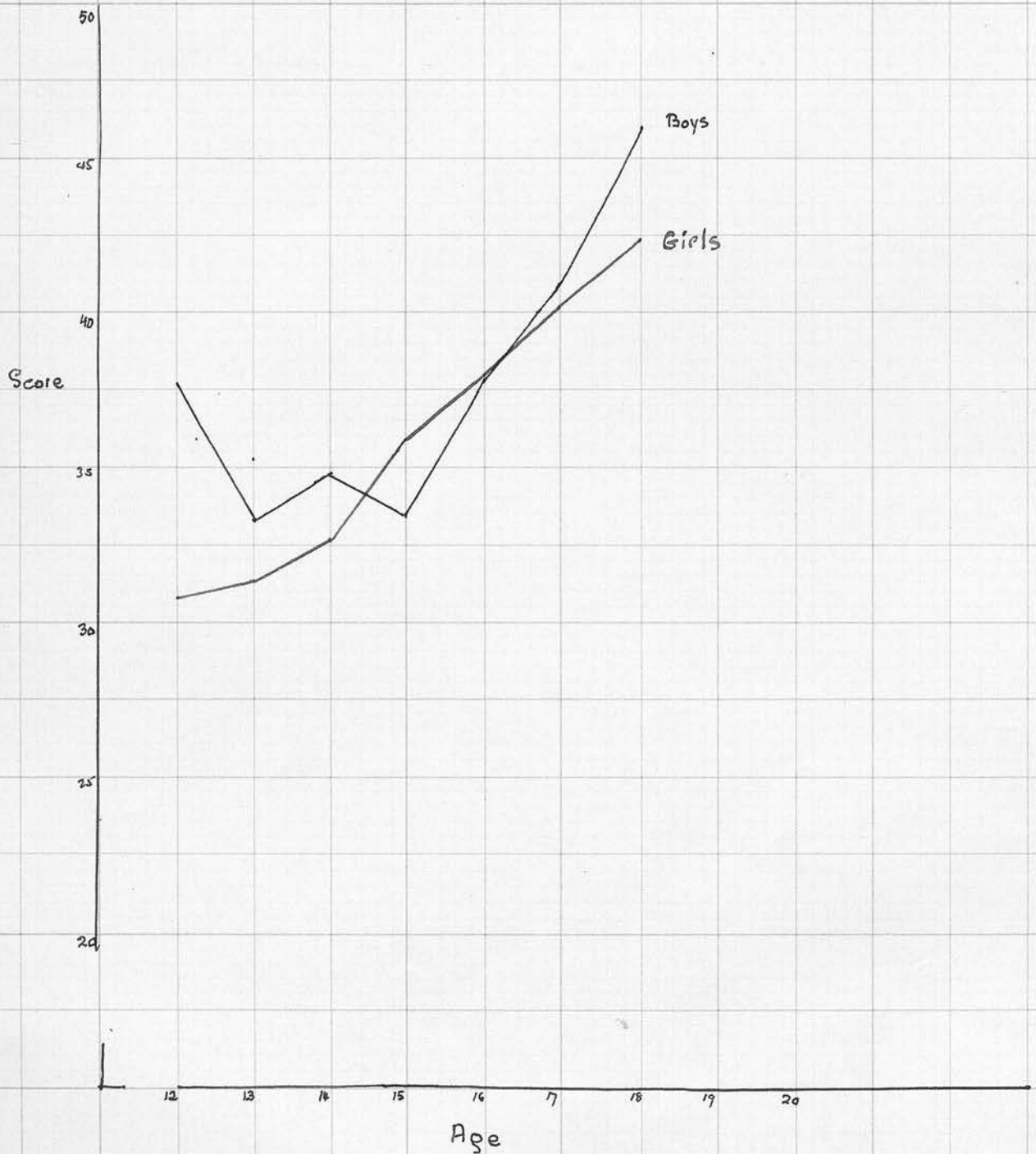
COMPARISON OF GEORGE WATSON'S BOYS' COLLEGE  
AND EDIN. LADIES' COLLEGE RESULTS.

---

Before/

# I<sup>st</sup> Test

Edinburgh Ladies' College & George Watson's Boys' College

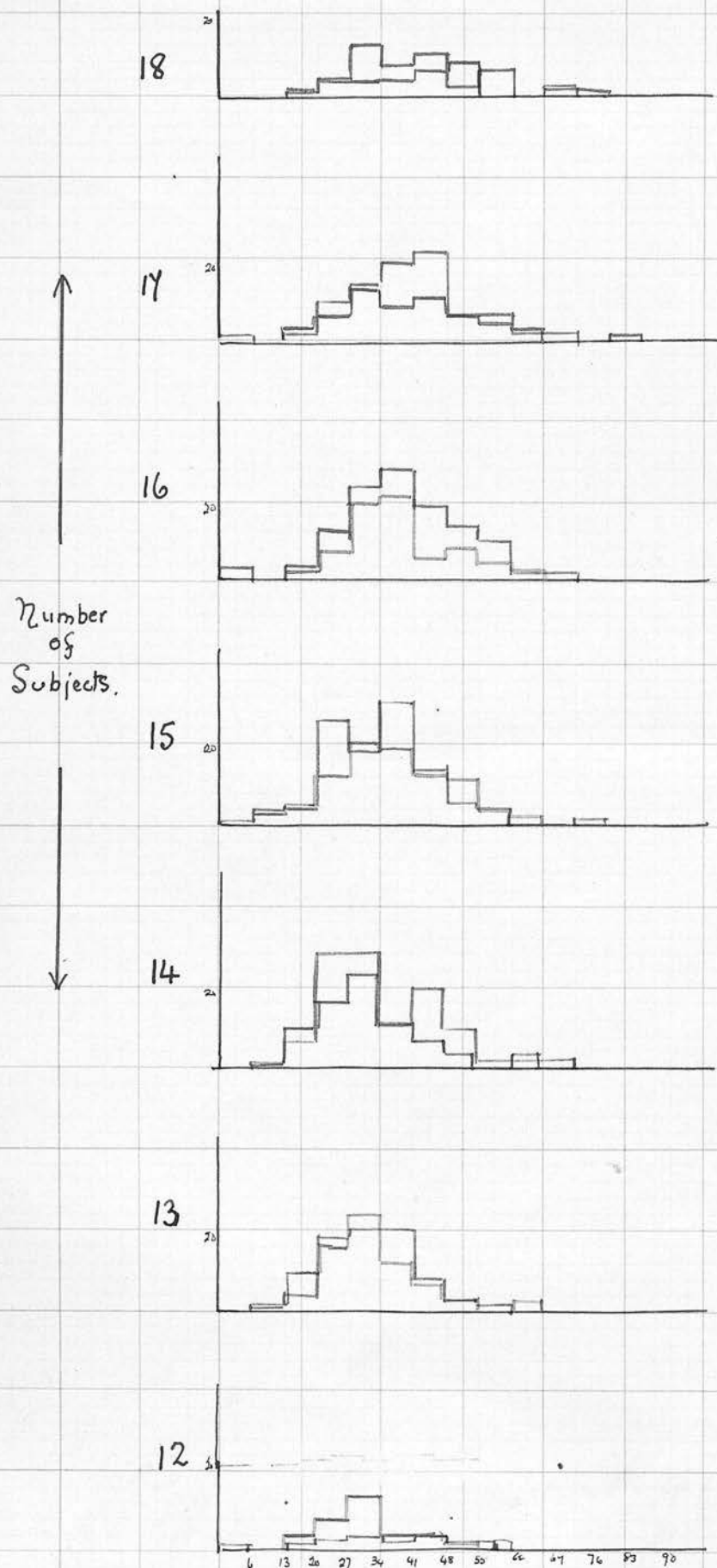


Scale. { 5 units horiz. = 1 class interval (7 units)  
2 units vert = 1 mark

Before any comparison is undertaken, let it be admitted that the results of the preliminary testing are too irregular to be entirely satisfactory. This holds for both boys' and girls' results. Since Edin. L.C. and G.W.B.C. are 'brother' and 'sister' schools a fair degree of similarity in the results may be expected. The expectation is not unfulfilled. Speaking generally, the schools are comparable throughout from the point of view of ability though the greater number of boys tested at each age hinders comparison to some extent. At certain ages differences in average ability are not unexpected. The greater variability of the boys usually remarked on by investigators is apparent up to 14 years, but not thereafter. The girls prove just as variable at the later years. This may be due to the vagaries of the distributions owing to the selection of ability at the later ages. It is of interest to note that as regards average ability the girls hold their own during the later school years. Superior intelligence is not generally regarded as being equally distributed in the sexes. A more rigorous selection may explain the equality evidenced in these groups.

The/

I<sup>st</sup> Test. G.W.B. College and Edin<sup>r</sup>. Ladies' College



Scale. { 2 units horiz. = 10  
5 units vert. = 20

{ Violet - Boys  
Green - Girls



The growth curves of boys and girls placed together (as on opposite page) reveal differences in average growth up to 15 years and an approximately similar pathway of advance from then to the end of the school period. That the dissimilarity before 15 years can be explained is shown in the previous analysis. The lesser difference after 16 years may be due largely to unequal <sup>numbers</sup> ~~members~~. The superimposed distributions of boys and girls on the following page serve to support the suggestions.

## ADULTS RESULTS.

---

In addition to the school children 345 adult subjects were tested in different groups. Of these 94 were men and 251 were women. They were drawn from colleges to which access could be obtained. Edinburgh University Psychology Class contributed 79 subjects (33 men and 46 women) between the ages of 19 and 24 years, Edinburgh Provincial Training College contributed 205 subjects, (44 men and 161 women). Heriot Watt College gave 17 men (engineering students) and Dunfermline Physical Training College 44 women.

The students of the Psychology Class were for the most part preparing for the degree of M.A. or already held it. The E.P.T.C. students were graduates in their year of training for the teaching of elementary or higher subjects, and ordinary College trained students in the two year course for non University men and women, which prepares chiefly for elementary teaching. The Heriot Watt College students were non-University men studying technical subjects. The Dunfermline Physical Training Centre trains girls to become Gymnastic and Games' Mistresses in schools and Colleges. Girls must be at least 18 years of age on entering. The course covers a period of/

## DETAILED AGES. ADULTS (344 Tested)

of two years.

The table below shows the distribution of adults, men and women, at each life age. The different sources of the subjects are also introduced. The sign 24 + is used to designate subjects of 24 years and above 24 years of age who are grouped together so as to constitute one age group after 23 years.

TABLE/

# EDINBURGH UNIVERSITY PSYCHOLOGY CLASS.

---

The students tested were 79 in number representing only half of the Psychology Class. The test was given by the experimenter. It has frequently been said that in testing adults the experimenter is in the hands of the subjects, whereas with children the reverse holds true. On this occasion the results prove that the subjects were not entirely serious as regards the performance. The directions were imperfectly read, though ample time for reading was allowed. Probably some of the subjects were quite familiar with what was required of them in a test. At all events, a number of them failed to observe that a number of questions were printed on the back page of the sheet. This was a misfortune. It appeared that the word "remaining" in the directions (referring to 3 pages) did not suggest to some subjects that a sheet has four sides. However, the fact was noted and never again was a test applied without the preface of a warning against this fault.

The results are briefly indicated below, but with due regard to the possibility of their unreliability. If this had occurred anywhere else than in the Preliminary Investigation all the results would have been



would have been discarded. But, at this point, their finality is not in question. Because of the small number of subjects at each age the results given refer to all the men, to all the women, and to all subjects taken together.

	Men	Women	All
Average Score	46	41.6	43.69
	33	46	79

The men's scores ranged from 15-77, The women's score from 22-77. There is little difference in this respect between the sexes, but the range for both is very wide.

By comparing the results obtained from this group with the results obtained from the schools, it is seen that the average adult score obtained from students whose ages stretched from 19 to 24 years and over is approximately the same as that obtained by 18 year old pupils. Were the results reliable and were the number tested large enough to make the score representative of the kind of subjects the indication would be that in the average students preparing for University/

University examinations in the Faculty of Arts have an average intelligence which is not superior to the average intelligence of 18 year old school pupils. But, in this case, both the qualifications are in question so that the exact placing of the group is not revealed.

#### EDINBURGH PROVINCIAL TRAINING COLLEGE.

A certain number of the subjects tested here were ordinary college students. These were tested in small class groups. The others were graduates and they were tested during an Education lecture period in one large mass. Practically all the men were University Graduates and about half of the women. More than half of the small number of men tested were 24 years and over. The girls ranged from 17 to 24 years with most cases between 18 ~~and~~ 21 years. The general results are shown below.

	Men	Women	All
Average Scores	50.7	42.7	44.5
St. Dev.	11.9	9.17	10.29
No. of Subjects.	44	161	205

It/

It is impossible to compare the men and women adequately on account of the totally different numbers tested in each sex. The superiority of the men can be explained by the fact that the girls average score is made up of two sections, a University and a hon. University one. The variability of the men equals that of the school children at the later years of school life; the variability of the women is noticeably small. The men's average exceeds anything found at the school period but the girls' average falls between the 17 year old and 18 year old age groups. A further diminution in the average score of the women is brought about by the elimination of the University women contributing to it. The average then becomes 40.01, that is to say approximately equal to the Edinburgh LADIES' COLLEGE 17 year old score.

The result obtained from this College goes to suggest that the averages found in the Psychology Class are probably lower than they should be.

The comparison of the results with the School results indicates that on the average the general ability of the subjects is level with that of 18 year old children.

HERIOT/

### HERIOT WATT COLLEGE.

---

The students at Heriot Watt College were tested by their lecturer in Engineering Mr MUIR. Mr MUIR followed the directions given him regarding the application of the test to the students. The number of students tested is small, only <sup>17</sup> ~~of~~ subjects between 19 and 24 + years being then available. The average age was 24 years exactly. The average score obtained was 42.35. The scores ranged from 16 to 61. These students would on the average fall into position between the 17 and 18 year old average scores obtained from the schools.

The group is undoubtedly selected and probably the men are generally superior in average ability. The small number of students tested is awkward for the formulating <sup>of</sup> definite conclusions. Extreme cases influence the averages. The median score is 45.

### DUNFERMLINE PHYSICAL TRAINING COLLEGE.

---

The test was applied by Miss YOUNG M.A., B, Ed., Assistant in the Dep. of Education at Edinburgh University and, ~~the~~ visiting Lecturer at the College. The girls tested were between 19 and 21 years (inclusive). Twenty were first year Students in the course. (aged/



(aged 19 and 20 years). Twenty-four were second year Students (aged 20 and 21 years).

The average number of marks obtained by the 1st Year group was 36.1 and by the 2nd year students 37.3. The average for all students was 36.79.

Again the number tested is too small for any great reliability to be placed in the results. Nevertheless the range of abilities, which is strangely small, indicates a homogeneity in the groups of subjects tested. In the first year group the marks ranged between 26 and 49, in the 2nd year group between 24 and 48. These students represent a level of development corresponding to that of 15 or 16 year old children. In reality the average score falls between these two years.



### COMBINED ADULT RESULTS.

---

So far, in the description of the results, each group of adults has been treated separately and the scores at each age massed together to give a grand total or average for all the men and women. When the groups are treated together it is possible to consider each age by itself and to compare it with other ages. In this way the character of the adult growth curve becomes clear. From the table previously given the number of adult subjects, men and women, at each life age is seen. Nearly three times as many women as men were tested owing to a difficulty in obtaining male subjects. The number of men tested at each life age is too small to be very useful, either for the purpose of obtaining an idea of the average ability of such subjects or for the purpose of comparison with the results of the women in the parallel age groups. With the women it is otherwise; and when the two are taken in conjunction matters are improved. About half of the men were over 24 years of age. The 24+ group is more satisfactory as regards numbers than the previous adult men's groups. At ages 19, 20 and 21, a fair number of women were tested but before and after these ages the numbers are again too small to be satisfactory/

satisfactory or representative.

It is evident in such a group as this that the subjects are selected. The adults tested cannot be judged representative of the average adult population. Perhaps they could not be further from representing it. It is rather true, that here is a fairly uniform group of "superior intelligence" subjects, those, as BALLARD would say, who have evaded the clutches of selective forces all along the line.

The table below gives the results obtained from the adult groups tested.

The distributions of scores from which the results are derived are placed beside the general discussion. They require no discussion. Their general inadequacy speaks for itself.

1st Test  
Adult Curves



Scale, { 5 units horiz. = 1 year  
2 units vert. = 1 mark

## ANALYSIS OF ADULT RESULTS.

---

The curves corresponding to the adult average scores are drawn opposite. As before the red line marks the average curve obtained from the combination of men and women's results.

### MEN.

---

With regard to the results obtained from testing men students it is evident that the numbers tested are too small to allow of adequate conclusions being drawn. The erratic turns of the curve prove the inadequacy of the testing. At 19 and 20 years the averages obtained are the same but at 21 years there is a marked descent in average scores. This descent proves temporary only since the curve again rises at 22 and 23 years. Up to 23 years averages are derived from scarcely a dozen subjects. Another drop in the curve appears from 23 and 24 years. The average score at 24 years derived from 46 individual scores carries more reliability with it than the adult mens' average scores at the other age groups. It is roughly equivalent to the 18 year old school score for both boys and girls and <sup>lower</sup> ~~more~~ than the average scores for men at any adult group. The 21 year/

# ADULT RESULTS.

AGE.	MEAN	P.E. of MEAN	St. DEV.	MEDIAN
19	M. (53 W. (40.4 ALL 42.12	(2.42 (.71 .76	(11.9 (8.75 10.22	(48.2 (41.46 42.8
20	M. (53 W. (40 ALL 41.49	(1.76 (.71 .75	(9.45 (9.03 10.36	(55.8 (39.7 40.7
21	M. (45.3 W. (42.65 ALL 43.14	(1.68 (.92 .83	(7.91 (9.1 9.1	(46.5 (41.47 42.2
22	M. (48.1 W. (43.94 ALL 44.88	(3.3 (1.17 .82	(15.89 (10.22 11.9	(52.6 (42.36 44.5
23	M. (50.2 W. (46. ALL 47.87	(2.3 (2.33 1.6	(11.3 (12.8 11.97	(53.5 (43.16 50.4
24	M. (44.63 W. (41.54 ALL 44.03	(1.54 (1.93 1.22	(14.28 (9.5 13.72	(44.1 (41.83 43.8

The adults falling into the 17 and 18 year groups are treated with the school children in the general results that follow. They are very few.



year old score comes nearest to it and since the subjects contributing to its make-up are <sup>from</sup> four different sources, it can be assumed not to be an extreme finding. Precisely at those points where University men preponderate are the scores highest. This occurs at 20, 22 and 23 years. The 19 year old score is made up from a few advanced George Watson's Boys College Results, and a few superior adult results.- hence its superiority. But that it is affected by extreme measures may be seen from observation of the median score at the same age. The variability of the groups calls for little comment. The differences in the standard deviations are largely due to chance selection which at some points has caused a "scatter" of scores and at others a decided concentration of scores.

#### WOMEN

The women's results are more satisfactory than the men's. At 19, 20, 21 and 22 years a fair number of subjects were tested and only beyond 22 years may great unreliability be expected. At 19 and 20 years the scores are almost entirely derived from Edinburgh Provincial Training College women and students of Dunfermline Physical Training College.

A lowering of the average score at these ages is not surprising since the Dunfermline Physical Training College girls have a lower average score than 16 year old school girls. The "cream" of intelligence is not concentrated in a Physical Training Centre. The students tend rather to be mediocre from this point of view with interests remote from academic studies. "Specific abilities" they may and probably do have. The adult curve is at its nadir point at 20 years where these subjects preponderate. From 20 to 23 years the average scores increase fairly steadily. After 20 years the average scores are obtained (with <sup>very</sup><sub>A</sub> few exceptions) from University women's test results. The topmost point of the curve occurs at 23 years. The high average score here may be due to the number of post graduate subjects contributing to it, but may also be influenced by the small number tested. At 24 years the curve again drops, as the men's curve dropped, but as the second column of the general table of results shows, no reliability can be attached to either this average score or that at the previous year. The variability of the women is nowhere so great as that of the men. No particular significance is attachable to this fact since the inadequate testing of men subjects accounts for the influence/

influence of the extreme measure found in their groups.

On the whole the average growth curve of the adult women is more satisfactory than the corresponding men's curve, but it is still far from being adequate. The erratic tendencies can mostly be explained as being due to the diversity or insufficiency of the subjects tested from age to age. It is noteworthy that the average scores at four different age groups lie approximately in a straight line. These are the average scores at 19, 20, 21 and 24 years. Only at 22 and 23 years when University post-graduates are in the majority does the curve depart from a fairly level course and at 22 years its departure is not significant.

A general comparison of the men's with the women's average curve points to a superiority of the men at every age beyond 18 years. The paths of the curves are altogether different in the first years beyond the school ages. This is due to the insufficient number of men tested, but also, and more significantly, to the difference in the sources whence the subjects come. From 21 years onwards the curves for men and women follow parallel courses but at different levels. Even the increase in average score at 23 years in the <sup>one</sup> ~~are~~ curve is comparable with/

with an increase at the same point in the other. Both men and women are University students at the ages, except for a few Heriot Watt College men entering chiefly at the later years. The average growth curve obtained by combining the men's with the women's scores does away with some of the irregularities, but owing to the preponderance of women at every age it conforms to the women's curve in shape, rather than to the men's. Since only 23 subjects go to produce the 23 year old score, it cannot be considered adequate. The reliability of the mean is proof of its insufficiency.

Viewed generally, the curve presents a "trough" and "crest" appearance on account of the low average scores of the 19 and 20 year old girls and on account of the high, inadequate scores of the 23 year old men. But apart from this it is not too much to suggest that, given a fairly uniform group of "superior intelligence" subjects the curve of growth or better, the curve of ability would present a flat appearance at approximately the 18 year old level, if a sufficient number of these subjects were tested.

The insufficiency of the data prevents further analysis.

TABLE CONTINUED

18	M. (46 W. (44.54 ALL 45.45	(1.21 (1.23 .86	(11.37 (11.3 11.3	(44 (45.3 44.8
19	M. (53 W. (40.4 ALL 42.12	(2.42 (.71 .76	(11.9 (8.75 10.22	(48.2 (41.46 42.8
20	M. (53 W. (40 ALL 41.49	(1.76 (.71 .75	(9.45 (9.03 10.36	(55.8 (39.7 40.7
21	M. (45.3 W. (42.65 ALL 43.14	(1.68 (.92 .83	(9.91 (9.1 9.1	(46.5 (41.47 42.2
22	M. (48.1 W. (43.94 ALL 44.88	(3.3 (1.17 .82	(15.89) (10.22 11.9	(52.6 (42.36 44.5
23	M. (50.2 W. (46 ALL 47.82	(2.3 (2.33 1.6	(11.3 (12.8 11.97	(53.5 (43.16 50.4
24	M. (44.63 W. (41.54 ALL 44.03	(1.54 (1.93 1.22	(14.28 (9.52 13.72	(44.1 (41.83 43.8



# GENERAL RESULTS of FIRST (or PRELIMINARY ) TESTS.

Since the Test was given with a fifteen minute's time limit to the girls of E. Ladies' G. to the boys of G.W.B. College and to all the adult subjects it is possible to combine these results to make a single curve for all the subjects. Unfortunately, the James Clark results must be excluded on account of the different time allowed for working the Test in that School.

The Table below gives the number of subjects at each life age contributing to the formation of the whole curve.

	12	13	14	15	16	17	18	19	20	21	22	23	24	TOTAL
MEN	20	74	113	105	109	80	40	11	13	10	10	10	46	641
WOMEN	34	72	85	34	69	51	38	70	74	44	34	13	11	579
ALL	54	146	198	189	178	131	78	81	87	54	44	23	57	1320

The table opposite gathers together the total results obtained from testing 1320 subjects. The continuity of development in the children's groups tested has already received attention. Comment on the adults group results followed. These facts need not be dwelt upon longer.

Test

Final Results

All

Men (red) Women (green)

24+

23

22

1

20

9

8

7

16

15

14

13

12

Number of Subjects

Scale { 3 units horiz. = 1 cl. interval  
5 units vert. = 10 subjects  
7 points  
Score

longer. It remains now to review the situation as revealed by the men's score, by the women's scores and by the two taken together.

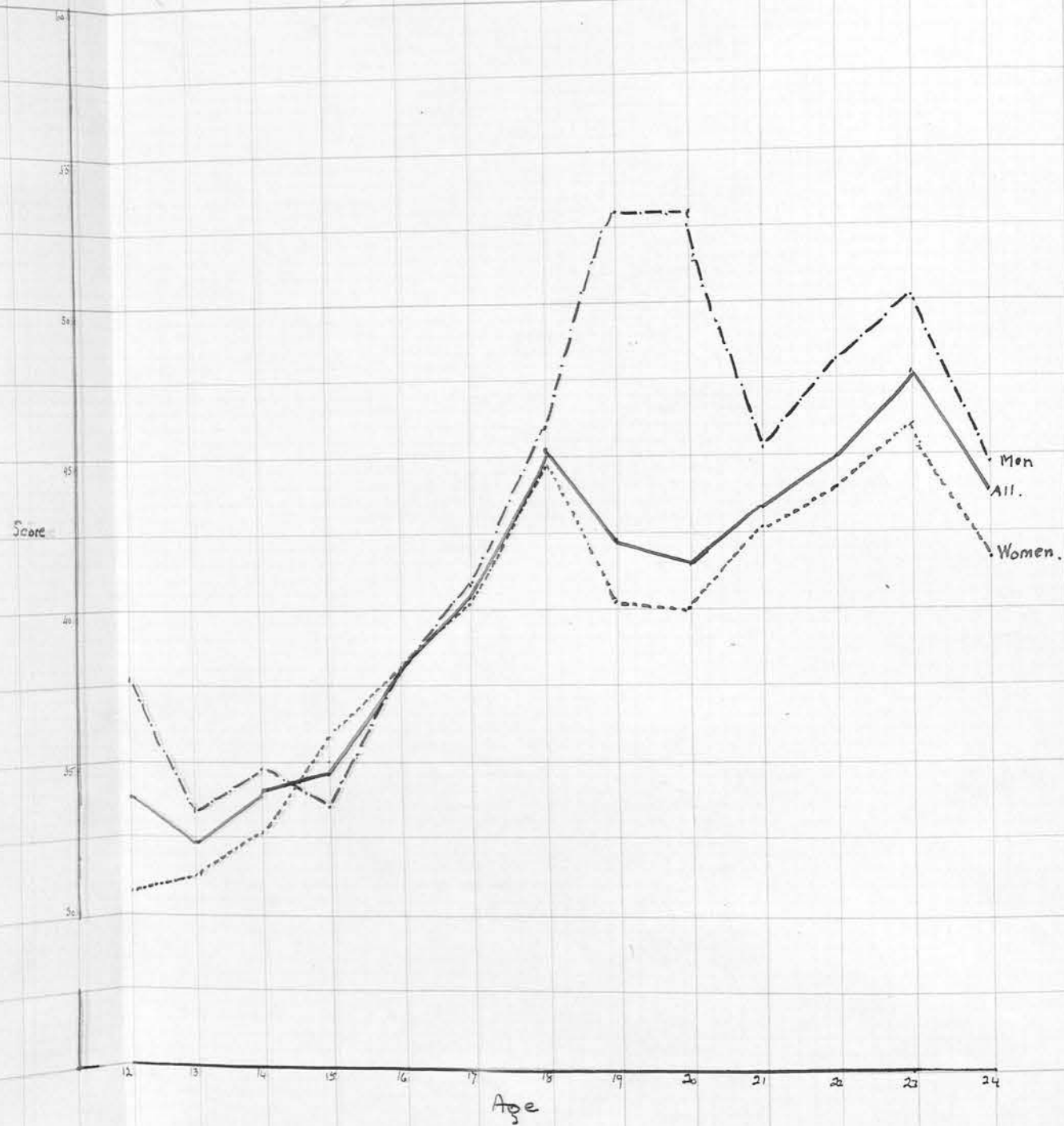
The pictorial representation of the situation as shown by the graph <sup>and distributions</sup> opposite may serve to make the facts clear. In interpreting it seems better to take the averages in conjunction with the growth curves which display them. In either case, it is unfortunate that the facts scarcely speak for themselves but require to be augmented or supplemented by additional information with as much veracity as possible. A too rigid adherence to figures profits little in this case, though as a basis for interpretation the figures are necessary.

#### MEN.

From the general chart of results certain things become apparent regarding the average ability of men as measured by the tests. As the picture stands, and without regard to anything else for the moment, there is a tendency for the average score of the men to increase but little from 13 to 14 years, to show little change between 14 and 15 years and after 15 years to show decided progress up to 20 years/

# I<sup>st</sup> Test

## Combined Results



Scale { 5 units horiz = 1 year.  
2 units vert = 1 mark



years, the average scores vary, but in no case do they exceed the 20 year old average. Rather, they tend generally to fall and at 24 years, after exhibiting erratic twists, the curve returns to the 18 year level. But the picture, as it stands requires to be supplemented. Up to about 15 years the variability of the groups increase. After 15 years it tends on the whole to be fairly constant except in cases where an inadequate number of subjects tested or a marked homogeneity in a group tends to increase or decrease the spread. The similarity of the groups from the side of variability renders comparison simple apart from other factors.

These other factors require careful consideration, for, taking the averages at their face value it would be easy to assume a gradual increase in average ability up to 18 years at least. The first difficulty would then arise when the need arose to explain the fall in the scores above 20 years. The complicating factors are due to selection, as has been seen. The unusual size of the average at 12 years coupled with an usually large variability measure has been explained as being due to selection, so, also the tendency of the scores to increase considerably/



considerably after 15 years. It follows, that the growth curve (opposite) is a misrepresentation of the truth regarding average mental growth, the subjects at 12 and after 15 being above average ability in general. The significant years then become 13, 14 and 15 years. In these groups average ability is presumably found.\* Moreover, it is in these groups that the flattening of the curve of growth was observed. Presumably at some point within the range or near its borders the upper limit of intelligence is to be found.

#### WOMEN.

If the women's average mental growth curve is free from the undue irregularities of the men's, it is still far from reaching perfection. Without adding to or commenting on the picture which the curve, derived from the average scores, presents, it can be seen that from 12 to 18 years there are indications of mental progress. After 18 years a decrease in scores, paralleled with a decline in the curve, is apparent. Only at 23 years is the 18 year old/

---

\* The question of Social status and its relation to ability must be ignored meantime.

old average slightly exceeded.

As in the case of the boys' curves additional facts must be noted in interpreting the curves. Up to 15 years selection does not operate significantly and up to 15 years there is still evidence of average growth. The variability of the groups is not so steady as in the boys' groups, tending now to exceed the boys' variability and now to fall below it. But, generally speaking, there is little difference except where differences may be expected. After 15 years growth seems to continue but the difference between 15 and 16 years is only slightly, if at all, attributable to pure growth, since selection commences to eliminate the poorer intelligences. After 16 years, even with selection operating the augmentation of scores is slight. These indications go to suggest once more that the upper limit of average mental growth enters somewhere near the fifteenth year but certainly not before then. The conclusion receives further support from the results of the Dunfermline Physical Training College girls. Although between 19 and 21 years of age, those subjects tested at a level between fifteen and sixteen years and of all adults tested these most nearly represent average/

average intelligence.

### ALL SUBJECTS.

---

Little remains to be said regarding the average growth curve which combines the results of the men and women and of the boys and girls. The averaging of the two tends to smooth away certain irregularities in the single curve. The interpretation of the curve is based on the same facts mentioned so often.

At the early years there is clear evidence of the flattening of the mental growth curve and beyond 15 years the course of development is obscured by the appearance of steady increase up to 18 years. The preponderance of E. Prov. T. C. and Dmf. Phys. TR. College, women at 19 and 20 years effects a lowering of the curve which rises after 20 years and remains approximately at the 18 years old level to the end.

The most significant feature of the general curve is the flattening between 13 and 15 years which goes to prove that average growth is then in the decline and cessation of growth must soon appear. The/

The increased scores at the higher ages, due to selection, indicate that, by the time the adult years are reached, ability such as is represented here, is that which falls, on the average, into the category "superior intelligence". It follows that the tendency for the older ages to approximate to the 18 years old level, (apparent at 21, 22 (23?) and 24 years) does not have a direct bearing on the upper limit of the growth of average intelligence as might be imagined. The turning of the curve at approximately 18 years is suggestive. It might on the one hand be assumed that the adult subjects, being selected, naturally fall beyond the average upper limit point. The theories of DOLL and BALLARD would countenance this view. But it seems that the turn of the curve at 18 years and the fairly constant approximations of the averages at later years permits of an alternative suggestion. It is not impossible, when the nature of the subjects at 18 years and beyond 18 years is remembered that, here is good evidence of an average upper limit for "superior intelligence". At 18 years, school boys and girls are assuredly, with few exceptions, of superior ability. The selective forces have weeded out all others. The same applies on the average to the University/

University subjects. All are not brilliant, but unless all are above the average intelligence of the people their success is limited in the academic Spheres. The homogeneity of the adult subjects tested here is fairly obvious. Where non-superiority of intelligence enters it has been duly observed, and it is where this non-superiority of intelligence occurs that the most significant deviations of the average curve from the 18 year old level are observed. Other deviations are due to chance selections based on inadequate numbers tested or to other explainable circumstances. The suggestion is only offered at this stage. Obviously, the inadequacy of the data does not permit of dogmatic assertions.

The conclusions arrived at in the preliminary investigations are stated with a full consciousness of insufficient proofs. At all points and with regard to all the groups tested, it is only possible to suggest that on the average, intelligence appears to grow up to 15 years at least (that is to say, to the  $14\frac{1}{2}$  -  $15\frac{1}{2}$  age group here). Beyond that point average growth is obscured by the entry of selective factors, but taking these factors into consideration it does not seem possible to place the upper limit of the growth of intelligence beyond 16 years or below 15 years of life age. Should it happen/



happen that further testing confirms this finding, the opinion of DOLL supporting an average upper limit of intelligence at  $13\frac{1}{2}$  years is once more to be rejected. As it is, it is difficult to accept his position.

In addition to the insufficiency of the data there may be added a possible inadequacy of the method employed. The study of averages obscures many things, though it is questionable if the obscuring is always a disadvantage. Nevertheless, it is true that a study of this type reveals next to nothing of the individual differences in ability which contribute to the whole. Further, the whole question of individual rate of development is entirely covered up. An investigation by means of individual testing extending over a large number of years is the most desirable method of approach to the problem, but it is not always practicable (as where continual access to schools is well-nigh impossible) and even when practicable, may be bristling with difficulties due to the time employed and more especially to the insufficiency of good testing material available.

A/

### A TIME LIMIT TEST.

---

The type of test employed here raises a question or two. It is a test with a time limit and it may be wondered if a 15 minute's paper is enough. Few investigations have been carried on to prove either the sufficiency or insufficiency of time-limit tests. In a recent Psychological Monograph (1924) BERNSTEIN reports on an investigation the purpose of which was to determine if there is such a thing as a speed factor in intelligence and if non-speed tests are of greater merit than speed tests. He finds no clear indication that subjects do better "in leisure" than in "haste" tests.

WOOLLEY and FISHER <sup>(1)</sup> report on their 1915 investigation with working children that the speed element in tests is not a great handicap, and that there is no discrepancy between estimated ability and rate tests. They rightly note that in the world the speed factor enters to a certain extent and in tests of "graded" difficulty it is also present only to a certain extent. Speed in their investigations was largely correlated with success in the tests themselves/

---

(1) WOOLLEY and FISHER. PSYCHOLOGICAL MONOGRAPHS.  
77.

LEIGH MUDGE'S (1) finding that speed tests are "rather accurate" and slow tests "rather more accurate" is faint commendation of the latter. FREEMAN (2) quotes an instance of the application of BURT'S Reasoning Tests in the Univ. of Chicago with a time limit to performance or no time limit to performance, which resulted in a correlation of zero. His own investigations lead him to conclude "It is conceivable that tests with a wide range of difficulties which are easy enough to permit those of little ability to make a score, but difficult enough to tax the ability of the ablest would give less correlation between speed and quality than a test which is easy and where speed alone counts". The matter is important since speed may handicap some individuals. It may be said in passing, that from tables constructed in the second part of the investigation, for another purpose, it was seen that there is evidence, even in as short a test as this, that correlation between Speed and intelligence is shown.

General opinion points to a sane conclusion that a "speed" test is not so reliable on the whole as a "leisure" group test or as a test which is a combination/

---

(1) LEIGH MUDGE: Journal of Educ. Psych. 1922.

(2) FREEMAN: Journal of Educ. Research. Jan. 1923

combination of both types. But, what is more important here is to know if a group test is as good as an individual test with older subjects, but on this matter nothing can be said definitely.

#### SEX DIFFERENCES.

Already some comparison between the sexes has entered incidentally into the general discussion leading to the belief that, generally speaking, sex difference in mental ability as measured by this test are lacking. An analysis of the scores obtained in the different tests contributing to the whole, might reveal sex differences, but in the preliminary investigations this analysis need not be undertaken, as it is not strictly relevant to the main problem. Suffice to say here, that where sex differences in the general scores are apparent they are due to factors of selection, complicating matters at a particular age group. In JAMES CLARK SCHOOL this occurs at 15 years. The difference between the boys and the girls there is 3.61 (P.E. 1.73) and is barely significant. In the other two schools it occurs at 12 years owing to selection, but at no stage thereafter with significance. The adult score, again, show marked/

marked differences but sufficient has been said to explain the occurrences.

The one point of interest regarding sex differences has already received comment. It is the flattening of the GEO. WATSON'S BOYS' COLLEGE average at 15 years. It suggests a natural "marking time" but is scarcely significant enough to be emphasised. Most writers note only slight sex differences as for example, YERKES. who sees difference only in regard to certain tests. But PORTEOUS in the article already spoken of (9. of App. Psych.) insists that girls are inferior to boys in general ability.

#### SOCIAL STATUS.

Since the first application of the BINET SCALE the question of the relation of Social Status to intelligence has been raised, but the complexity of the problem and the "infancy" stage of mental testing hinders the resolution of the difficulty. DECROTY and DEGAUD, two Belgian experimenters with the first BINET TESTS found  $1\frac{1}{2}$  mental years of difference between children of superior social Status and children coming from humble homes. BINET himself slightly/



slightly experimented and proved the truth of their results. Since then nearly every significant experiment has contributed some evidence to support the early findings. In his early experiments, BURT found indications of the influence of social status on test results. Also in his more recent work, <sup>(1)</sup> he compares two schools "typifying the best and worst in a representative borough" and finds that the superior school is nearly a year ahead of the general average. Regarding the differences found in the report by early investigations, <sup>(2)</sup> STERN has this to say. "One might assume that the very rise to the higher and better-off classes would itself predicate a certain intellectual selection and that the children of these classes would have come into the world equipped with a superior intellectual endowment. BURT, remembers that tests test endowment in conjunction with all the influences to which the examinee has been subjected up to the moment of testing and it is just these external influences that are different in the lower social classes". This touches once more the assumption of a common environment which the tests presuppose.

STERN'S/

- 
- (1) BURT: Mental and Scholastic Tests. p. 191  
 (2) STERN. Psych. Method of Testing Intelligence.

STERN'S own investigation with the BINET tests applied to Volksschule and Vorschule children resulted in the finding that the normal 10 year olds in the former (and inferior) type of school had an average intelligence equivalent to that of the normal 9 year olds in the latter (and superior) type of school.

The results of the testing undertaken here bear out the conclusions of most writers that the tests are not uninfluenced by social status. The JAMES CLARK SCHOOL results indicate decidedly lower average scores than the other schools at the lowest ages but the difference becomes less marked towards the end of their school career. Doubtless, selection operating at the higher ages in the JAMES CLARK SCHOOL explains the approach of the 96 average scores to the average scores of the G.W.Bs and E.L.C. children at about 15 years of age.

#### INTELLIGENCE and SCHOOL SUCCESS.

The results of the tests were sent to the schools tested in order that an opinion of their worth might be given. No definite assessment of intelligence/

intelligence was asked from the schools. The Headmasters of G.W.B.C. and J.C. School found some discrepancy between their estimates of the children's intelligence, derived from several sources, and the test results. Similarly the score sheet was passed round the entire staff of Edinburgh Ladies College. The staff expressed interest in if not approval of the undertaking. Comments were not always favourable but, by the courtesy of Miss Mary G. Clarke, then Headmistress of the School, the investigator was given <sup>confidential</sup> access to the record sheets of school marks at the end of the year. The correlation obtained between test scores and school work was .70!. BURT quotes a correlation of .91 between his revision of the BINET-SIMON Scale and school work (p. 182 op.cit) so that there is every reason to be content with the results obtained from this short and probably inadequate time limit test.

## SECOND INVESTIGATION.

---

December 1924 - January 1925

### Subjects.

The second part of the investigation is virtually an attempt to test the reliability of the preliminary testing. In order to do this it was decided to retest the schools used in the first investigation, and, access to these schools being again granted, the children of the James Clark School, both boys and girls, the boys of G.W.B. College and the girls of E.L. College were tested in Dec. 1924 and Jan. 1925. The sources from which the adult subjects were obtained were, the Psychology Class of Edinburgh University, the Education Class of Edinburgh University, various classes in Education and Psychology at Edin. Prov. Tr. College and three Technical Engineering Classes at the Heriot Watt College. Only a part of the Psychology Class was tested, as the remainder were members of the Educ. Class. It was possible to effect a separation. In the table below the numbers from each source are shown.

SUBJECT/

## SUBJECT

## 2nd. INVESTIGATION

School or College	Men	Women
G.W.B.C.	316	.
E.L.C.	.	383
J.C. School	284	303
Ed. Univ. Psych. Class.	14	43
" " Educ. "	50	138
Heriot Watt College.	59	.
Edin. Prov. Tr. College.	9	107
Total	1032	974
All	2006	

The Tests.

The retesting of the same groups as were examined before fitted in with the idea of constructing a new test, on general lines similar to the first but with a slightly longer time limit. In the construction of the new test certain alterations were made in the choice of material. Mirror reading was left out. It offered too much temptation to children to/



to children to hold the sheet up against the light. The test which was termed somewhat inadequately "word penetration" suffered the same fate. The previous results had shown that the Logical Reasoning test was decidedly too easy and could be answered perfectly by nearly all the young children. Another, more difficult, might have been found, but it was preferred to introduce some new material. The last test, comprehension of a passage, was also left out.

The new test sheet, again constructed on the 'cycle' principle contained four types of test which were separated four times. These were analogies, directions, number completion series, and proverbs. They constituted 16 out of the total 20 test questions. The remaining four were two orientation tests, one set of 12 hard opposites and finally, a code test. Each set of analogies and numbers contained six parts and three problems were set in each proverbs test. All the numbers series problems were originally constructed on the old principle (also the proverbs tests, the directions and orientation tests.) As before, the gaps in the number series completion tests were placed at any part of the line. The opposites test words were called from several sources in the Journal of Educational/

Educational Psychology. Only what were considered difficult words were chosen. The code test was formed after a model in the Simplex Group Test prepared by Richardson.

Directions were placed on the front page of the test sheet and spaces for the subject's responses placed alongside the questions. This plan, adopted in box tests, is a labour saving device both for the subject and the examiner. Responses took the form of a letter, a word or a figure to be written in all but one case. The code test alone required a little writing.

#### Time Limit - Tests. Key to Correct Solutions.

##### Marking:

The method of marking again was somewhat arbitrary and similar to the method adopted before. For each simple analogy or each number series completed, one mark was given at all stages of the paper. Each section of a proverbs test correctly answered obtained two marks, i.e. a total of six marks for each of tests 5, 10, 15 and 19. Directions and orientation test received 4 marks each if correct and no marks if wrong. In marking the code test, the method adopted was to give three marks for the grasp of the principle and an additional mark for each word correctly/

correctly written. The total was 12 marks. The sum of all the marks was 120 but for practical convenience all the total scores were turned into percentages.

TIME LIMIT. The test was very carefully tried out on children and adults of superior intelligence in order that a suitable time limit might be affixed, with absolute certainty that none should succeed in finishing all tests in the given time. Twenty minutes seemed to fit the requirements.

PSYCHOLOGICAL EXAMINATION.

Time Limit: \_\_\_\_\_ Minutes.

Name \_\_\_\_\_

College \_\_\_\_\_

Date \_\_\_\_\_

Age \_\_\_\_\_

Do not open this pamphlet until you are told to do so by the Examiner.

On the inside pages and on the back page there are a number of problems. In each case you are told exactly what to do. Read the instructions carefully.

Do not ask any questions. If you come to a problem that you do not understand, go on to the next problem.

You will be given 20 minutes. Solve as many problems as you can in the time allowed.

Solve the problems in the order given. Begin at the beginning, and go straight through.

DO NOT TURN THIS PAGE UNTIL YOU ARE TOLD TO DO SO.

PSYCHOLOGICAL EXAMINATION.

- (1) Each of the lines below, when completed will/

will contain four words, the first two of which are in the same relation to each other as the second two. In each case there is one term missing. Fill in the blank with a word.

# EXAMPLES.

(a) finger is to hand as toe is to  
(The missing word is "foot".)

((b) cabbage is to as cod is to fish.  
(The missing word is "vegetable".)

sky is to blue as grass is to .

day is to as white is to black.

is to navy as soldier to army.

pupil is to as child is to parent.

go is to come as is to buy.

fin is to fish as wing is to .

(2) Notice these two numbers, 3, 5. If there are as many days in September as in March write the larger number here , but if the reverse is the case, write the smaller number here .

(3) Fill in the missing steps in the following number series:-

# EXAMPLES.

(a) 2 4 6 8 10 12 .. .. (The missing numbers are 14 16.)

(b) /

(b) 1 5 2 5 .. 5 4 .. (The missing numbers are 3 5).

10 9 8 7 6 5 .. ..

8 1 6 1 4 1 .. ..

10 .. .. 25 30 35 40 45

27 27 23 .. 19 19 .. 15

6 12 .. 24 30 .. 42 48

22 19 16 .. 10 .. 4 1

(4) The streets of the town run north, south east, and west. I set out for a walk, and after walking straight ahead for some time, I turn to the right. A few minutes later I turn again, this time to the left, walk for a considerable distance, and then turn once more to the left. When I set out I was facing south. In which direction am I now facing?

(5)

1. Every rose hath its thorn.
  2. He who would eat the kernel must crack the nut.
  3. Every shoe fits not every foot.
  4. As you make your bed, so must you lie.
  5. Every herring must hang by its own gill.
  6. No autumn fruit without spring blossoms.
- Which/



Which of the sayings above means the same as  
the following:-

- (a) As you sow you must mow. Number
- (b) Every path hath its puddle. Number
- (c) Paddle your own canoe. Number

(6) Fill in the blanks in the following:-

cellar is to attic as bottom is to .  
eye is to head as window is to .  
is to laughter as sorrow is to joy.  
skin is to body as is to tree.  
behind is to late as is to early.  
cold is to hot as Pole is to .

(7) If the letters in the word "begin" occur in  
the same order in the word as they do in  
the alphabet, and if this is true of the  
letters in the word "finish", write "K"  
but if it is only true of one of the words,  
write the last letter of the word.

(8) Fill in the missing steps in the following  
number series:-

16	12	..	11	14	..	13	9
2	7	4	..	..	11	8	13
3	6	8	16	..	36	..	74
...	..	43	35	27	19	11	3
5	8	10	..	15	18	..	23
..	14	21	..	35	42	49	56

(9) Fill in the blank space next each word  
with the word which means the opposite.

friend	animated	serious
former	venturesome	sinful
robust	clumsy	genuine
deceitful	haughty	pride

- (10)
1. Little by little the bird builds its nest.  
speak
  2. All are not friends that one fair.
  3. One should not always judge by appearances.  
A
  4. There are no gains without pains.
  5. He that hath no money needs no purse.
  6. Like master, like man.

Which of the says above means the same as the  
following:-

(a) All are not white who are millers,  
nor all black who dig coals.

Number .

(b)/

(b) He that hath some land must have  
some labour. Number .

(c) Little strokes fell great oaks.

Number .

(11) Fill in the blanks in the following:-

dismal is to dark as cheerful is to .

iron is to as cambric is to fabric.

horse is to as docile is to stubborn.

fire is to smoke as water is to .

complex is to feminine as is to masculine.

food is to body as is to engine.

(12) In the following list of words draw circles

round three words so that two are included,

one of which is the opposite of the

other, and the third means the same as

one of these two-

wet stiff dry straight moist

(13) Fill in the missing steps in the following

number series:-

7	3	..	3	5	..	4	3
11	14	..	16	15	..	17	20
51	50	47	..	43	..	39	38
1	4	..	..	22	25	50	53
20	16	18	14	31	34	14	10
23	26	..	..	31	34	35	38

(14) A/

(14) A certain building has windows on two sides only - on the side directly opposite the door, and on the side which is to the left of the door as you enter. The early morning sun shines straight through the door. In which direction (north, south, east, or west) does that side of the building face which has neither window nor door? Answer .

- (15)
1. Perseverance overcomes all difficulties.
  2. Plants oft removed never thrive.
  3. Pour not water on a drowned mouse.
  4. He who is in the mud likes to pull another into it.
  5. Joan is as good as my lady in the dark.
  6. Unhappy experiences teach us to be careful.

Which of the sayings above means the same as the following:-

- (a) Do not add affliction to misery.  
Number .
- (b) When candles are out, all cats are grey.  
Number .
- (c) Failure follows frequent change of plan.  
Number .

(16) Fill/

(16) Fill in the blanks in the following:-

motive is to          as why is to how.

failure is to success as depressed is  
to          .

truth is to          as beauty is to art.

copy is to originate as imitate is  
to          .

politeness is to          as manners is to  
morals.

character is to reputation as truth is  
to          .

(17) Write in the space below the first letter  
that in this sentence goes immediately be-  
fore the letter which in the alphabet  
goes immediately after it. Answer          .

(18) Fill in the missing steps in the following  
number series:-

..	8	10	..	22	44	46	92
143	114	88	..	..	28	14	3
29	28	..	25	..	23	21	20
63	..	31	..	15	14	7	6
20	17	..	14	11	9	..	5
..	14	17	..	16	12	15	11

(19)/



- (19)
1. A tree is known by its fruit.
  2. Waste not, want not.
  3. Wood that grows crooked will never be straightened.
  4. Don't cry over spilt milk.
  5. Don't cross a bridge till you come to it.
  6. Gather ye rosebuds while ye may.

Which of the sayings above means the same as the following:-

- (a) Take Time by the forelock.  
Number .
- (b) What is bred in the bone will never out of the flesh.  
Number .
- (c) Never meet trouble half-way.  
Number .

- (20)
- A message is to be sent in a "code" or secret language. First the whole message is written backwards (in capitals), and then there is substituted for each letter the letter which comes next but one before it in the alphabet.

EXAMPLE - He is coming: first written backward, thus GNIMOCSEH.

When for each letter is substituted the next but one before it in the alphabet, the message becomes ELGKMA QG CF.

Now/

Now put the following message into code form.  
Write only the final form. Use capitals all the  
time:—

Enemy entrenched five miles to south-east. Rein-  
forcements following.

Print code here.

# KEY to CORRECT ANSWERS.

1. green: night: sailor: ~~teacher~~ or master: sell: bird:
2. 3.
3. 4,3: 2,1: 15,20 23,15: 18,36: 13,7.
4. East.
5. 4; 1: 5
6. top: room: tears: ~~bark~~: before, in front: Pole or  
Equator.
7. N.
8. 15,10: 9,6: 18,37: 59,51: 13,20: 7,28:
9. foe or enemy: latter: sickly: open, straightfor-  
ward:  
lifeless: cautious, careful: agile, graceful:  
humble: frivolous: sinless, holy. Godly.  
counterfeit spurious: humility:
10. 3; 4; 1.
11. bright or light: metal: mule: steam: simple:  
fuel, coal:
12. wet, /

12. wet, dry, moist.
13. 6,3,: 13,18: 46,42: 8,11: 16,12: 27,30:
14. North.
15. 3: 5: 2.
16. Method: elated: science: invent: behaviour: be-  
lief or opinion.
17. e in the.
18. 4,20: 65,45: 26,24: 62,30: 15,8: 18,13.
19. 6, 3, 5.
20. ELGOMJJMD QRLCKCAPMDLGEP RQYC-FRSMQ MR QCJGK  
CTGD BCFALCPRLC WKCLC.

### The Application

The tests were applied to the children in G.W.B.C. and E.L.C. on the same day in Dec.1924. Dr. Drever and Dr. May Collins tested the boys in George Watson's Boys College, the investigator the girls in Edinburgh Ladies' College. The boys were once more grouped in two very large groups in the exam hall. The girls were tested in their classrooms. In George Watson's Boys College the testing covered little over an hour, in Edinburgh Ladies College, it required about three hours. The James Clark children were tested in January of this year. The adult groups were tested by the investigator at dates/

dates more or less convenient for the lecturers. Mr Muir tested two of the three Heriot Watt groups after having seen the application of the test to the third group.

Procedure.

The papers were distributed with the injunction that they were not to be turned over until a direction to that effect was issued. A few minutes were devoted to filling in the details regarding name, data etc. and to the reading of the directions. The subjects were given no further information. They were told all to observe carefully the signal to start and the signal to stop.

## JAMES CLARK SCHOOL.

---

The data chosen for the application of the test was perhaps a little unfortunate here. The school organisation was suffering from a recent great influx of pupils (mostly girls) from another school, and also from the presence of painters in certain parts. The facility with which the testing was carried out in the difficult circumstances was largely due to the industry and help of Mr. Ambrose, Senior Assistant and head English Master in the school. By reason of his interest the whole school was tested in two days.

On account of the recent influx of pupils (they had been in the school only one week) the number of children attending the school was considerably augmented. 587 children were tested. Of these 284 were boys and 303 were girls. The sexes were thus fairly equally divided. In the table below the numbers between 12 and 16 years are given.

As before 12 means  $11\frac{1}{2}$ - $12\frac{1}{2}$  years and so on.

Table/



## BOYS.

AGE	MEAN	P.E.OF MEAN	ST.DEV <sup>n</sup> .	MEDIAN
12	22.01	.72	7.91	21.86
13	21.95	.56	8.75	21.73
14	29.84	.83	11.76	29.86
15	35.38	1.64	12.39	33.5
16	Too few.			

## GIRLS.

AGE	MEAN	P.E.OF MEAN	ST.DEV <sup>n</sup> .	MEDIAN
12	21.66	.68	8.19	20.68
13	24.7	.56	9.1	24.27
14	26.80	.87	12.39	26.5
15	37.66	1.74	11.77	37.68
16	Too few.			

## ALL.

AGE	MEAN	P.E.OF MEAN	ST.DEV <sup>n</sup> .	MEDIAN
12	21.81	.49	8.05	21.75
13	23.40	.39	8.75	23.7
14	28.31	.60	12.18	27.91
15	36.61	.85	11.9	36.29
16	( 33.22	2.8	13.3	37.5 )

T A B L E.

	12	13	14	15	16	TOTAL
Boys	54	108	91	26	5	284
Girls	65	120	93	21	4	303
All	119	228	184	47	9	587

There is inequality of numbers from age to age, as was found in the previous investigation, but there is approximate equality in the number of each sex at the different age groups. The 12 year old children are probably not selected; - , the 16 year old children probably 'old dull' pupils. The small number at 15 years clearly indicates the operation of selection after 14 years, and shows how few entering the school at the beginning of the course manage to reach the end. As stated in the last investigation, it is probable that selection on the basis of ability begins only at fifteen years and not earlier.

The results obtained from the school are placed/



← Score →

Scale. { 3 units horiz. = 1 c. interval (7 marks)  
 { 5 units vert. = 10 subjects.

placed below. On the opposite page the distributions from which these average scores were obtained are illustrated.

From the results embodied in the tables the average growth curve from 12 to 16 years is constructed.

### ANALYSIS OF RESULTS.

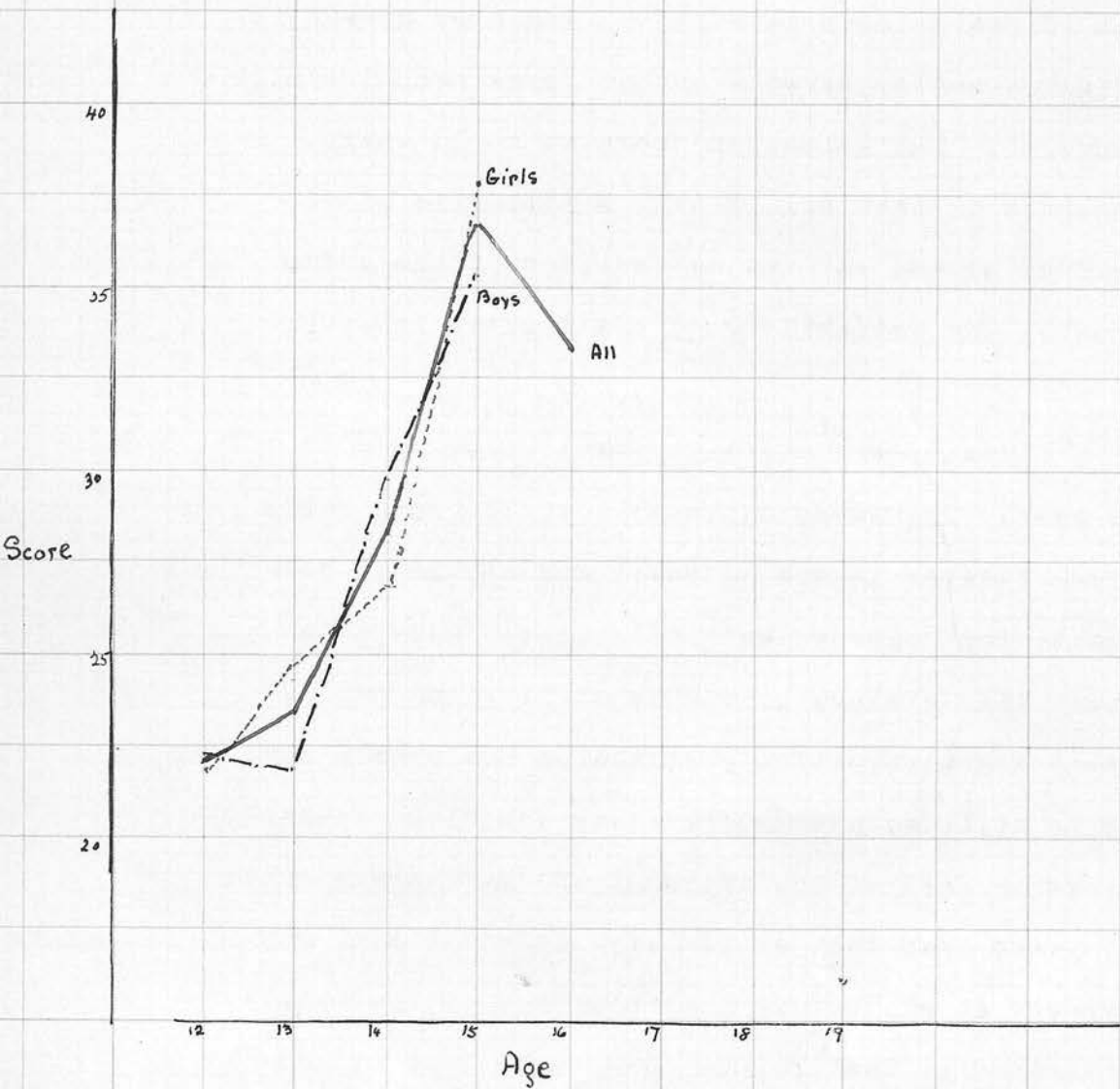
---

#### BOYS

The results are remarkably similar to those obtained in the preliminary test. The 12 year old boys slightly surpass the 13 year old average. The lower age group contains brighter children on the average and the number only half as many as there are at 13 years. From 13 to 14 years there is indication of considerable average growth coupled with increase in variability. That the growth is greater than the average growth indicated in the previous test may be due merely to a difference in units in the two tests, but it is more likely to be due to the influx of 'low' subjects at 13 years drawing down the average score. The skewness of the 13 year old distribution may be explained on this ground also. The previous year's distribution at 13 years is not noticeably/

## 2<sup>nd</sup> Test

### James Clark School



Scale 5 units horiz. = 1 year  
2 units vert. = 1 mark.



noticeably skewed. The influx affects not only the 13 year old group but tends to pass over into the group beyond though to a much less extent. Between 14 and 15 years there is still evidence of growth if the figures and appearance of the curve are literally interpreted. But selection operates at 15 years. The average at that age is only a composite of 26 individual scores and the variability of the group is great. The reliability of the figures is doubtful.

The growth curve, somewhat flat from 12 to 13 years, increases noticeably at 14 years. The interval between 13 and 14 years may appear to be increased for reasons mentioned before, but it is apparent that average growth is still going on. At 15 years where selection complicates the growth curve there is still an average increase visible. Probably the greater part of the increase of the average score at 15 years over that at 14 years is not due to mental growth at all but to the selection of the brightest children at that point. The curve does not extend far enough to suggest the real point at which average mental growth ceases. The indications are that it is not sooner than 14 years at least

The/

The table showing the percentage of each group reaching or exceeding the median of the group below ought to give a general idea of the change from one age to another, but the figures again require to be supplemented by other information.

---

Percentage of 13 year old reaching or surpassing the median of the 12 year olds	51
Percentage of 14 year olds reaching or surpassing the median of the 13 year olds	74.7
Percentage of 15 year olds reaching or surpassing the median of the 14 year olds	65

---

At 16 - too few cases.

---

The figures at 13 years falsely suggests that average mental growth is on the point of reaching its maximum. The figure at 14 years falsely suggests that average mental growth is not becoming less. This high percentage finds a counterpart in the average for 13 years which has already been discussed. The table of percentages, on the whole, requires that too much allowance be made for one thing and another not apparent on its surface representation.

#### GIRLS.

In order to facilitate comparison the distributions of the girls' scores are placed beside the/  
the/

the corresponding series for boys. The girls' average at 12 years is not so high as that obtained by the boys of the same age, probably on account of the fact that a larger influx of girls than boys occurred. The result of the influx is a counteraction of the selection which normally is characteristic of the 12 year old children in the school. The effect of the influx of pupils is seen to a larger extent in the 13 year old and 14 year old age groups. The fact that twice as many 13 year olds as 12 year olds were tested, must be taken into consideration in observing the 13 year old average. But, regarding the 12 year old average as a fairly reliable indication of the average mental ability of the 12 year old girls, the average score at 13 years may be taken to show a real increment. The difference between the 12 year old mean and the 13 years old mean is probably not as great as it would be, if the presence of a large number of strange children at 13 years did not tend to lower the average.

The distributions of scores at the point and at the following year exhibit a slightly skewed appearance in the negative direction. The skewness is, in the girls' distribution at 13 years less/

than  
 marked in the corresponding boys' distribution but it  
 is most noticeable at 14 years. At 14 years, again,  
 the average score shows an increase, which may be  
 taken as real, despite a small number of children  
 pulling the average down. The distribution clearly  
 shows the influence of the lower scores and the  
 variability measure is unduly large on their account.  
 At 15 years the average increase in score is too  
 large to be reliable. Also, it is based on few test  
 scores. The reliability of the mean is in question  
 and the variability measure large. The increase at  
 this point corresponds to that shown by the boys.

On the whole the evidence is still fairly  
 inconclusive. The influx of new pupils is a disturb-  
 ing element, interfering with the homogeneity of the  
 groups. The figures up to 14 years, and it may be  
 assumed that they are reliable indications of what is  
 actually going on, point to a fairly steady but small  
 increase in average score. Beyond 14 years the course  
 of development is quite obscured. The short curve  
 of growth shows that the average growth which is so  
 slight up to 14 years takes a new course thereafter,  
 which can only be allowed for when the operation of  
 selection /

selection is noted. The curve and the scores (when due allowance for variability differences is made) indicate that soon after 14 years, other things being equal, the average growth curve must become flat. A further and a conclusive support to this suggestion comes from the table showing the percentage differences from year to year which is placed below.

---

Percentage of 13 year olds reaching or surpassing the median of the 12 year olds	65.8
Percentage of 14 year olds reaching or surpassing the median of the 13 year olds	52
Percentage of 15 year olds reaching or surpassing the median of the 14 year olds	85
too few	

---

At 14 years the distribution only succeeds in surpassing the 13 year old group by a very small amount. Presumably, by 15 years, the real difference would become negligible.

A general comparison of boys' with girls' results shows, as before, that the boys tend to be somewhat superior to the girls. The extreme scores may not be true representations of difference, in which/



which case the general sex differences in mental ability are slight. The distributions do not tend here, as before, to lend support to the view that the variability of boys is greater than that of girls. But the evidence here is inconclusive.

#### COMBINED RESULTS.

---

The combination of boys' and girls' average scores shows that the situation is favourably comparable with the results of the preliminary investigation. The same slight increase of average score from 12 to 13 years is apparent; there is a considerable increase from 13 to 14 years; there is a marked increase from 14 to 15 years due to the operation of selection, and a marked decrease at 16 years due to a superabundance of dull pupils and a lack of subjects in general. The distributions strengthen the conviction that the influx of new or strange pupils causes a skewness in the curves at 13 and 14 years and it may be to some slight extent at 12 years also. The successive groups show an increase in variability which is erratic at points. The great variability of the girls at 14 years is carried over to the general/

general distribution, to break the continuity in variability which might be expected. The small range of years and the disturbing elements at nearly every stage go to lower the expectation of perfect results in this respect.

The form of the distributions, incomplete though these distributions are, suggests that at 13 years there is a continuance of average mental development, which, being very small, probably indicates a decrease in average growth. The course of development in the average or combined curve, is obscured from this point onwards. The explanation is to be found in the increase in the average score made by the boys of 14 years over their average score at 13 years, the score at the last mentioned age being itself diminished. The combined score, therefore, does not speak for itself. A truer representation of average growth seems to be available in the girls' results. Beyond 14 years no reliability can be attached to the results. They are insufficient in themselves.

The average mental growth curve itself exhibits a changed course after 13 years on account of the complications. Obviously it falsely represents the real course of development. It is not likely that/

that an average slight increase between the younger ages 12 and 13 years (granting the superiority of the former average and the inferiority of the latter) should be offset by a great increase in average growth at a later age.

The addition of the percentage differences, bearing on the change exhibited by the groups taken in their entirety, reveals next to nothing of the true state of affairs. It requires no comment!

---

Percentage of 13 year olds reaching or surpassing the median of the 12 year olds	58
Percentage of 14 year olds reaching or surpassing the median of the 13 year olds	60
Percentage of 15 year olds reaching or surpassing the median of the 14 year olds	72.3
Percentage of 16 year olds reaching or surpassing the median of the 15 year olds	44

---

It seems that, in this case, the combination of scores scarcely proves advantageous. It covers up what it is desired to show. At the same time it proves that no support can be lent to the view that intelligence ceases to grow at  $13\frac{1}{2}$  years life age. The opinion expressed before is reiterated, that up to 15 years of age there is good reason to believe that average mental/

mental growth continues and that after that point growth is either greatly minimised or absent altogether.

It will be observed that, in the course of the discussion, it has been indirectly implied that the JAMES CLARK children are of superior intelligence to those who entered the school under unusual circumstances. This is believed to be true. The study of the scores of these children adds the proof. The scores are extremely low in certain cases. But the implication needs further comment. It might be, and probably is, generally assumed that the JAMES CLARK school recruits, not the average child, but children slightly above the average. That, though generally true, cannot be supported. The difference exhibited between these school children and the other school children is explainable at the lower years by other means. The other school was of less than average ability. It reveals the same characteristics as contrasted with the JAMES CLARK school, as were found in previous significant studies of social status differences. The knowledge of the fact that very many/

many children at the JAMES CLARK school come there only to occupy their last two years of compulsory school attendance is some proof that they are not selected. That which selects them, and which selects always in such cases, is parental goodwill, good sense, or the material possibility of transfer to such a school. Individual cases differ but the general rule remains. The value of technical education has hardly yet made its appeal to the home environment except in a desultory fashion. This is not a denunciation of the school. Far from denunciation it merits the most praise that can be given it. The writer firmly believes in its worth. It receives children from the elementary schools who are strangely dependent on others. Independence they have not. The curious thing is that this school teaches them what they lacked. The difference at the early stages of the school and at the later was so apparent to the writer that it called for comment. The answer was that the children were taught "to think for themselves". The exact bearing of the by-talk on the question at issue may be wondered. It has been indulged in to point out that some part of the ability//



ability shown by children at the later ages may be due to the good training received in the school. In the manner of going about the test the children may differ, and that the manner or method of attack makes a difference cannot be denied. It does not seem desirable that investigation should shirk the revelation of any factors whatever.

To return to the topic of discussion, and in conclusion, it may be said that the idea of selection on the basis of ability choosing the children at the JAMES CLARK School cannot be countenanced. Teachers with whom the matter has been discussed agree that, where theoretically selected ability is catered for, practically all and sundry enter.

This is the justification of testing the school to obtain the upper limit of the growth of intelligence. Apart from the assumption further allowance would be necessary for the fact that the samples of children were not random. It does not appear to be necessary.

# GEORGE WATSON'S BOYS COLLEGE.

---

The boys were tested in two large groups. The twenty minutes test was applied to the first of these groups by Dr. Drever and to the second group by Dr. May Collins. More boys were tested this year than last by the inclusion of some classes at the lower end, comparable with the lowest classes tested in Edinburgh Ladies' College in the first test. Unfortunately, the investigation was forced to omit these very classes in retesting Edinburgh Ladies' College. In all 616 boys were tested.

The table below shows the number of pupils tested at each age.

14 =  $13\frac{1}{2}$  -  $14\frac{1}{2}$  years etc.

AGE.	12	13	14	15	16	17	18	19	TOTAL
NO.	53	87	128	129	110	80	25	4	616

At 18 years the numbers are so small as to make the results of questionable worth. At 12 years the subjects are probably not so selected as on the previous investigation in which only those advanced for their years were included. That selection operates/

2<sup>nd</sup> Test  
George Watson's Boys' College

↑  
number  
of  
Subjects.

18

17

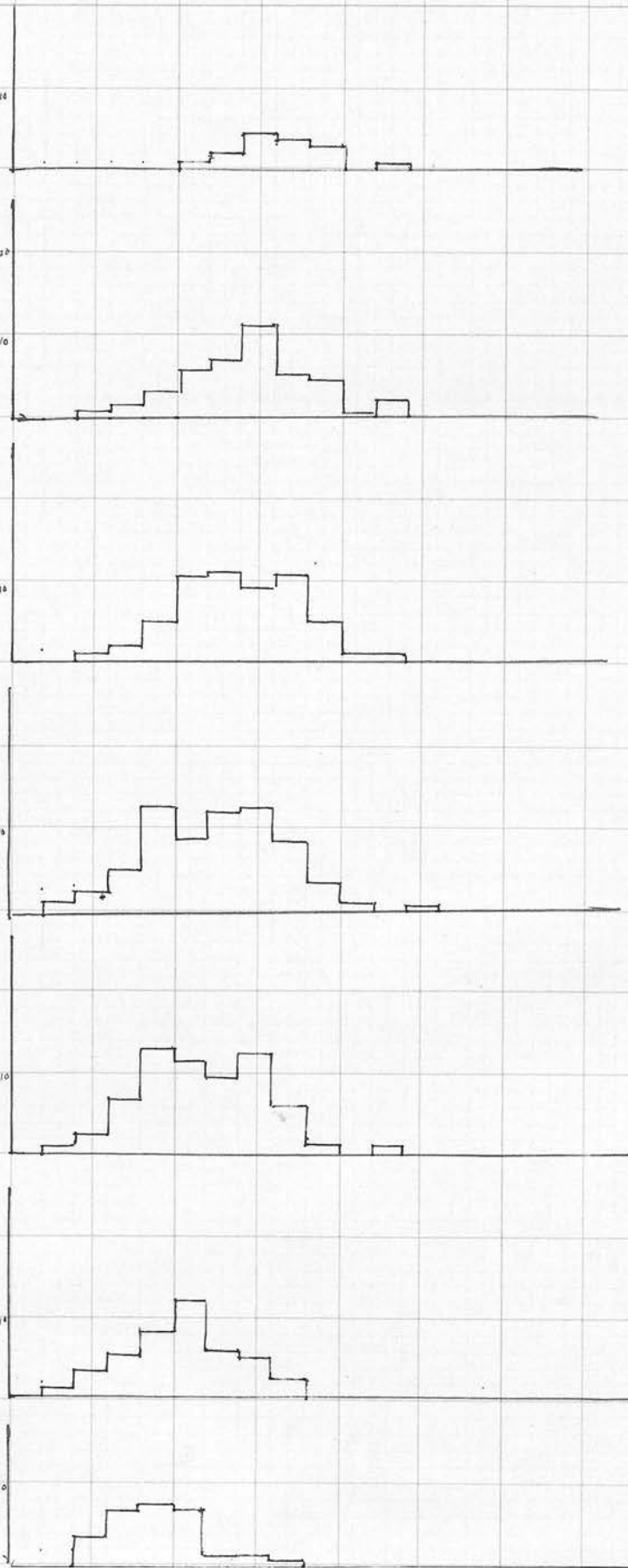
16

15

14

13

12



← Score →

Scale. 2 units horiz = 10

5 units vert = 10

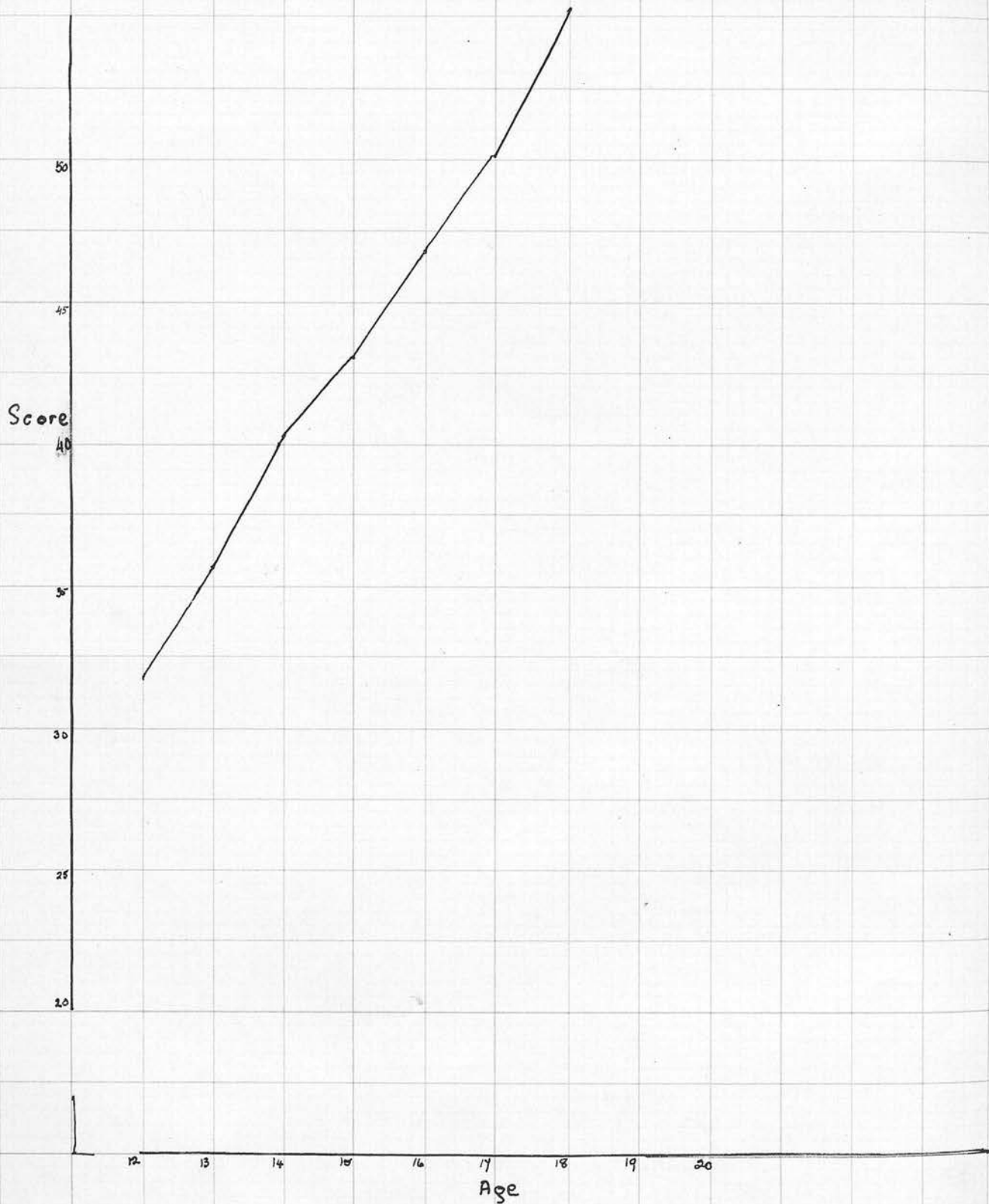
operates only after 15 years has been verified by observation of the scores of those pupils tested in the first year who have not returned to school. There is only the comparison of two years' results to show this, it is true, but the testimony is otherwise supported. At 15 years the poorer intelligences begin to drop out. Beyond 15 years the groups become selected and interpretation must take devious routes.

The distributions of scores are placed opposite. General observation shows them to be fairly normal and to indicate average increase from 13 to 17 years inclusive.

The results derived from the distribution are given in the table below. The curves obtained from the average scores are drawn on the following page.

Table/

Test.  
George Watson's Boys' College



Scale.  $\begin{cases} 5 \text{ units horiz} = 1 \text{ year} \\ 2 \text{ units vert} = 1 \text{ mark. (point)} \end{cases}$



AGE.	MEAN	P.E. of MEAN	ST. DEV.	MEDIAN
12	31.60	.88	9.59	31.42
13	35.62	.82	11.4	36.98
14	40.31	.74	12.39	40.08
15	43	.74	12.6	43.45
16	46.89	.77	12.09	47.66
17	50.11	1.01	13.44	51.1
18	55.24	1.37	10.15	56.56

It is well to remember that in the second test a better representation of the 12 year old ability was obtained than in the preliminary investigation. Consequently, it is not surprising that the results indicate a greater difference between the average scores at 12 and 13 years than was found before. This more representative value at 12 years makes clear that from 12 to 13 years and from 13 to 14 years average growth is still progressing. The two intervals indicate approximately the same amount of growth. It is probable that whatever influx of scholars occurs at these two age groups does not influence the distribution one way or another on account of its particular nature. Selection operating at 15 years makes the average difference between/

between 15 and 16 years slightly greater than it was before. This probably indicates that average mental growth is diminishing because to attribute as much of the difference to actual mental growth as is observed between 12 and 13 years and again between 13 and 14 years would be to leave a negligibly small part of the whole difference to explain a strong selective influence. At 17 and 18 years each advance over the previous score again shows the work of selective influences. At 18 years it is doubtful if such weight can be attached to the value given, owing to the small number of subjects tested. Beyond 15 years it is difficult to trace the true growth curve.

The variability of the groups increase from 12 to 15 years but thereafter, <sup>probably</sup> on account of the selected groups it becomes erratic and at 18 years it is noticeably low. On the whole, however, the groups are comparable generally. The distributions indicate a slight increase of average growth at least to 16 years. After 16 years of age no general progress is obvious.

Returning now to the actual figures, it appears that average mental growth changes are apparent/

apparent up to 15 years when selection enters. The average difference between 15 and 16 years is not appreciably greater than the difference found at the previous age, as has been seen. It seems, then, justifiable to assume that average mental growth progresses a little from 12 to 15 years and after 15 years becomes less apparent. The table of difference adds confirmation to the view that by 15 years or near 15 years the summit of mental growth is reached. The table is quoted below.

DIFFERENCES IN AVERAGE SCORE FROM YEAR TO YEAR.

INTERVAL	DIFFERENCE	P.E. OF DIFFERENCE
13 yrs over 12 yrs	4.02	1.2
14 " " 13 "	4.79	1.1
15 " " 14 "	2.47	1.4
16 " " 15 "	3.89	1.06
17 " " 16 "	3.22	1.24
18 " " 17 "	5.13	1.7

Beyond 15 years the differences are of no importance. The P.E. of a difference based as it is on the assumption of a random sampling, is useless thereby itself. Besides the differences existing beyond 15 years are known to be due to selection. But as/

as the table stands, it proves the reliability of the differences up to 14 years - that is to say, during the first two intervals above 12 years. At the third interval (15 years over 14 years) the difference is only doubtfully significant. (The difference is 2.47 and is less than 3 times the P.E. 1.06) The conclusion seems to be that after 14 years the average mental growth is negligible or only very slight - probably the latter.

The same conclusion results from the study of the percentage differences from age to age which take the whole distributions into consideration. The actual figures are such as to support the view that mental growth continues right up to 15 years. The selection entering the field at 15 only serves to keep the difference between the 15 and 16 age groups equal to the difference between the periods two. This indicates at least a diminution in growth after 15 years. The table below contains the percentages discussed.

Table/

	Percentage of 13 year olds reaching or surpassing the median of the 12 yr. olds	
"	" 14 "	68.9
"	" 15 "	62.5
"	" 16 "	58.1
"	" 17 "	58.7
"	" 18 "	57.5
		72



This diminution after 15 years is itself only the second in the series, as can be seen from the table of average differences from year to year. Between 14 and 15 years the average increase is less than in the two earlier age groups. Consequently, it is safe to assume that the mental growth decreases which commences to show itself after 14 years will continue and zero growth become apparent after 15 years. Now it is inconsistent with the theory held that the George Watson's Boys College pupils should on the average continue to develop in mental ability for a slightly longer period than the James Clark Children also taken on the average.

The curve of average mental growth rather conceals then reveals the situation. But, up to 14 years it indicates a steady increase in average mental growth and from 14 to 15 years indicates the diminution already mentioned. The results certainly discredit DOLL'S  $13\frac{1}{2}$  years level if they do not justify the fixing of the upper limit of mental growth at any definite point in the age series.

# EDINBURGH LADIES COLLEGE.

The testing was carried on by the investigator according to the method already indicated. Fewer girls were tested with the second than with the first test. The 'lower' 12 year olds were omitted as it was inconvenient to have them, and one group of 'poor' 15 and 16 year olds (to 16½ years approximately) was regrettably but unavoidably left out. In all 383 pupils received the test. The distribution of the total number at the different ages is indicated below.

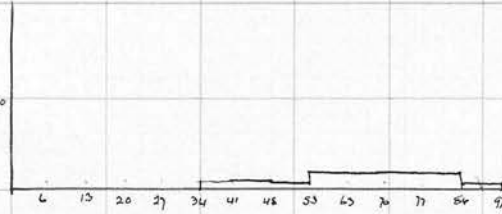
AGE.	12	13	14	15	16	17	18	19	TOTAL
NO.	13	67	81	76	64	53	21	8	383

In comparison with the same table for G.W.B.C. the numbers here are small. At the extremes of the scale the numbers tested are again small and at the two age groups 15 and 16 they are less than in the previous test, owing to the number omitted. Comparison of the second test results with the first seems to indicate that selection as the basis of ability/

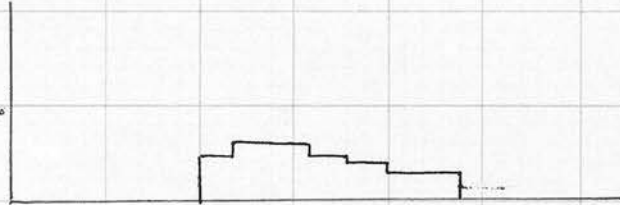
# 2<sup>nd</sup> Test

## Edinburgh Ladies' College.

18



17



number  
of  
Subjects

↑  
↓

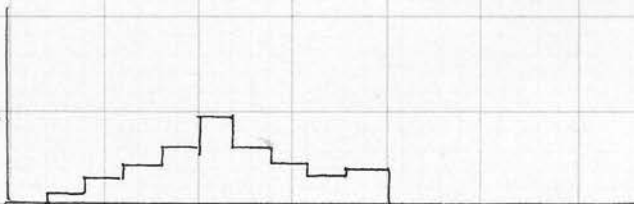
16



15



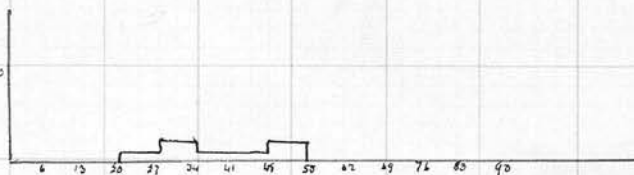
14



13



12



Scale { 2 units horiz. = 10  
5 units. vert. = 10

← Score →

ability commences slightly earlier with the girls than with the boys. That is to say at 15 years there is evidence of its working and although the tendency is not strong it requires to be taken into consideration.

The distribution of test scores are placed opposite. Owing to the smaller numbers tested at each age, they present a more flattened appearance than the corresponding boys' distributions. It was considered useless to change the limit for mere exhibition purposes. The distributions are only of interest as the basis of all calculations and to show the general movement at each age group. Up to 16 yrs. there is evidence of a continuance of mental growth. After that age small numbers prevent the observation of progress, if progress exists. The distributions, however, are not all-revealing.

In the table below the results of the testing are shown-:

Table/

AGE.	MEAN	P.E. of MEAN	ST.DEV. <sup>n</sup>	MEDIAN.
12	39.53	1.92	10.43	41.25
13	37.32 (36.56)	1.02 (.89)	12.96 (10.92)	35.8
14	40.38	.99	13.3	39.5
15	44.15	1.06	13.7	44.75
16	52	1.23	14.7	54.45
17	54.18	1.11	12.04	53.85
18	59	1.9	13.3	59.62
19	.	too few.		

At first sight the table is not at all reassuring and if no explanation could do away with some of the curious results the position would be far from satisfactory. Three things are most striking in the general picture; the size of the means at 16 and 18 years, the extraordinary and changeable variability measures, and the values of the P.E. of the medians at 13 and 16 years in particular. A comparison of these test results with those of G.W.B.C. suggests at once that some things are wrong. The girls' average is much superior to that of the boys at 16 years, it shows, in fact, an inclination to superiority all through, the variability measures are greater and as a/



2<sup>nd</sup> Test  
Edinburgh Ladies' College

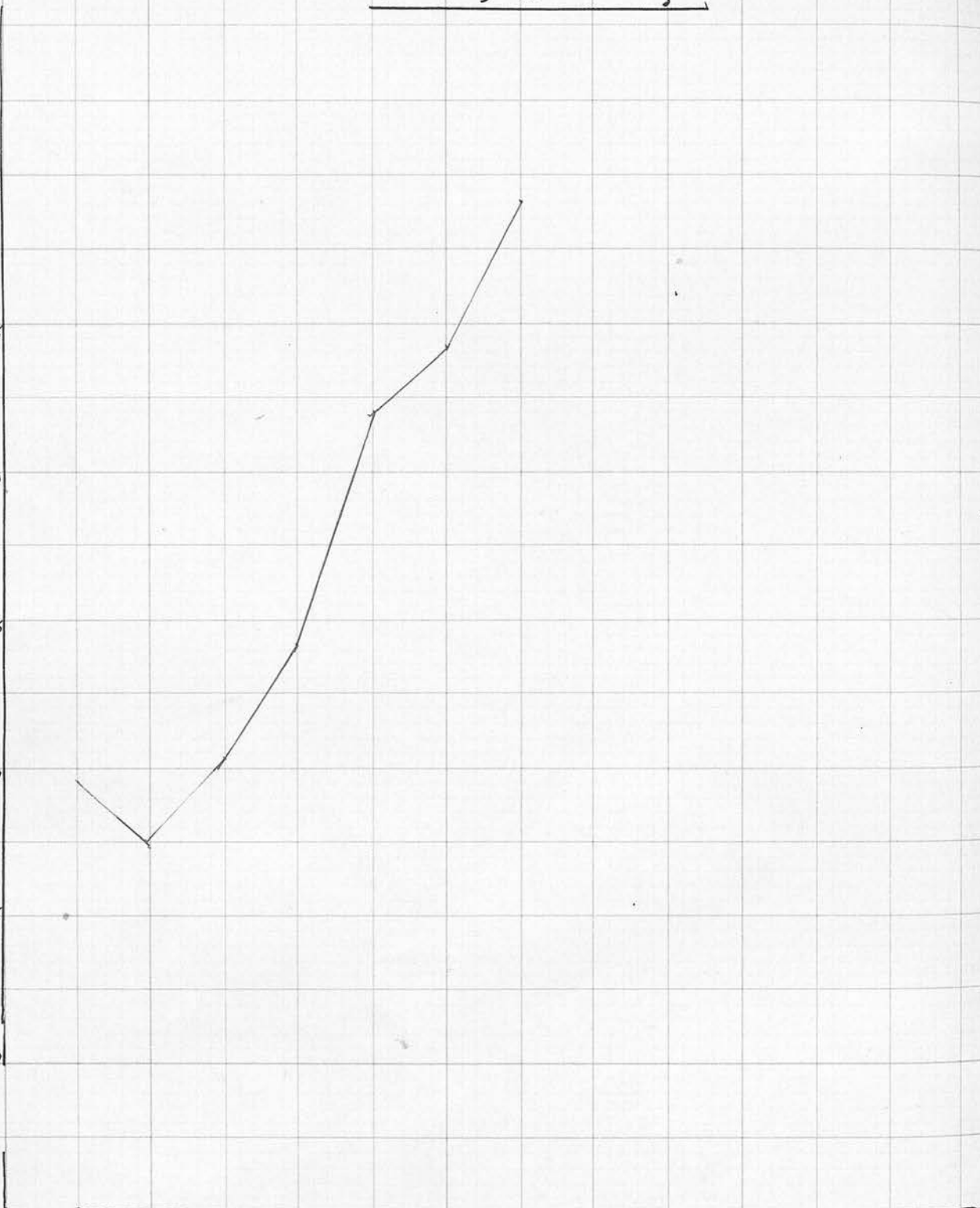
Score

60  
55  
50  
45  
40  
35  
30

12 13 14 15 16 17 18 19 20 21

Age

Scale. { 5 units horiz = 1 year.  
2 units vert = 1 mark



a consequence of the increased variability and the smaller numbers tested at each age the P.E. of the mean assumes rather large proportions at times. A closer analysis of the situation is demanded, though in general and apart from its superiority over the boys' curve of growth, the girls' curve (placed opposite) exhibits the tendencies noted in the previous investigations both curve and table show a heightened average mental ability at 12 years, due to the selected material tested at that age. From 13 to 14 years there is evidence of an increase in mental growth which is repeated in the following interval. The second interval is, curiously, slightly larger than the first. Between 15 and 16 years the average increase is unexpectedly large. A slight increase occurs at 17 years and again at eighteen years the difference increases. Beyond 15 years there is no doubt that selection must go to augment the score. At the later ages the groups are not representative of average intelligence. At 16 years some explanation of the large average scores must be forthcoming and doubtless the explanation will extend to make clear the cause of the large variability measure at the same age. The variability is larger than at any other age in the table and the P.E. of the mean points to a degree of unreliability. Any explanation/

explanation based on a difference in the application of the test to this group is impossible since the group is made up of elements from several classes. The cause of the trouble lies in the unfortunate omission of 'poor' subjects excluded from the testings at this age. The variability measure, comes from a distribution which nearly presents a 'chimney pot' appearance through the absence of subjects who would have gone to build it up at the lower end.

The situation at 13 years also demands explanation. At this age the girls surpass the boys and exhibit a distinctly greater variability. This fact defied analysis till inspection of the distribution shows that the results are influenced at this age, by the presence of one abnormally bright subject. Leaving this subject out of account the figures summarising the situation are reduced (They are extended in brackets below the actual figures derived from the distribution. The mean becomes 36.36, the standard deviation 10.92 and the P.E. of the mean .85).

One other fact demands comment. The variability of the groups in the girls' school is strangely greater than in the boys' school. In large measure the comparison is futile since the numbers tested are/

are totally different in the two schools. The difference in the standard deviation at 15 years, (13.7 for the girls and only 12.6 for the boys) is not so baffling when it is observed that 76 subjects contribute to the girls' deviation and 129 to the boys'. The erratic movement of the girls' standard deviation is another matter, and these have been explained above.

The striking peculiarity of the results explained, it remains to attempt their interpretation. The table of average differences from age to age is placed below though it does not facilitate analysis. The many distributing factors in the results necessitate much additional by-talk. The P.E. of a difference method is scarcely legitimate beyond 15 years where the groups are selected or at 12 years for the same reason.

INTERVAL.		DIFFERENCE.	P.E. of DIFF.
13 yrs. over 12 yrs.		-	-
14 " " 13 "		3.06 (3.82)	1.4 (1.33)
15 " " 14 "		3.77	1.4
16 " " 15 "		7.85	1.6
17 " " 16 "		2.18	1.65
18 " " 17 "		4.82	1.9

The bracketed results refer to the group with the unusually bright subject omitted very little profit is to be derived from the study of the figures in the table.

The average at 12 years requires no further comment. It explains the descent of the curve at 13 years. Between 13 and 14 years there is evidence of some increase in average mental growth and the same holds good between 14 and 15 years. (The 'Difference Tests' barely admits the reliability of the difference. The P.E.'s are scarcely three times the corresponding difference). In regard to average growth the girls' results are comparable with the boys at the same age. The average growth that is evident is not large. There is, on the contrary, good reason to believe that it is diminishing. Working from the actual average obtained at 13 years this is less apparent than working from the reduced average, but if this procedure is objectionable it does not need to be resorted to. For, it is probable that selection enters in a little before  $15\frac{1}{2}$  (or before the beginning of the 16 year age group. If selection plays any part in the  $14\frac{1}{2}$  to  $15\frac{1}{2}$  age group/



group (i.e. 15 years) then the real average difference between the age groups 14 and 15 will rather be less than greater than the average difference between the previous age groups 13 and 14. This would point to a diminution of average growth and an approach to the upper limit of the growth of intelligence. One further fact requires to be noted. The omission of one group of 'poor' selection affect not only the 16 year old group but also the 15 year old group. That is to say, there is at least some possibility that the average score found at 15 years is slightly in advance of what it would be if this group had been included. On the other hand, it might be true that while such subjects would affect the higher age group for the worse they might not similarly affect the lower.

Therefore, as the figures stand, they require one or other of the explanations offered above, since an increased average difference at the later age is contrary to the Laws of growth.

Strictly/

Strictly speaking the figures do not allow any fixed point indicating a cessation of average mental growth, to be named. The possibility is, that they do not contradict the findings in the other groups. One last avenue of enquiry is open in the study of the percentage increase of each group over the preceding group.

PERCENTAGE  
of

13 year olds reaching or exceeding)	12 yr. olds	3.4
the median of the )		
14	13	64.2
15	14	61.8
16	15	68.7
17	16	54.3
18	17	61.9
19	18	.

If these percentages indicate the general progress of the entire groups they may be interpreted to mean that average mental growth increases definitely up to 15 years, and is not to be found after 16 years. The 16 year old percentage is irregular, as would be expected, and conceals the exact stage at which average mental growth ceases, but the indications are that it must be above 15 years or rather above the  $14\frac{1}{2}$  to  $15\frac{1}{2}$  age group.

The truth is, that the results are not straightforward enough to be satisfactory. The writer is convinced of the veracity of the 'interpretation', but others may not accept it willingly. The unsatisfactoriness is not without interest. It evokes many questions, such as the following:- Does it make any difference whether subjects are tested in large or small groups? The G.W.B.C. results are almost as perfect as could be expected in both years of testing. Does it make any difference if subjects are aware that they are doing a test for "one of themselves"? Or generally, does the personality of the tester count for anything? Be it confessed that the Edinburgh Ladies' College girls in the later stage groups, by whom the experimenter was most likely to be remembered, exhibited marked 'zeal' in responding to the test. BURT, in the early days commented on the influence of the same factor on test results. Any individual child may be zealous but, when a few are so, the zeal is contagious. This factor, however, is not assumed to enter into test situations any more than into other situations. Now is not the time to answer these questions.

Moreover, taking into consideration the unsatisfactory nature of the Edinburgh Ladies' College/  
 lege/

College results, it does not seem necessary to add a general comparison of this school with George Watson's Boys College. That the two resemble one another in reality, as was found in the preliminary testing, is without doubt.

## ADULT SUBJECTS.

### Second Investigation.

The number of adults tested in the second part of the investigation was 420, an increase over the numbers in the preliminary testing. Of these subjects 288 were women and 132 men. They were obtained from the different Colleges already mentioned in the introduction. The total number of subjects from the different sources is given in the table below. The unequal representation of the sexes at each group is evident.

	MEN	WOMEN
Edin. University Psychology Class	14	43
Edinburgh Prov. Tr. College (2 year course)	9	107
Heriot Watt College	59	
Edin. Univ. Education Class	50	138
TOTAL	132	288

The results from the different colleges are first of all tested separately and in a general manner.

EDINBURGH/



# EDINBURGH UNIVERSITY PSYCHOLOGY CLASS.

---

The section of the Psychology Class tested was selected to avoid testing a certain number of subjects who would also be likely to be members of the Education Class, since they were aiming at obtaining a University Diploma in Education. In all probability the best individuals were not included in the group tested since the Diploma Class has a large representation of post-graduate students who take the Diploma Course during their period of training for the teaching profession.

This year the subjects were well warned against failing to notice the fourth page on the test sheet, so that on that account no trouble could arise. The sexes are represented rather unequally in this group, there being only 14 men against 43 women. The men's results cannot on this account be considered very reliable. As the general table (included later) shows the ages ranged from 19-24+ years.

The results obtained are as follows:-

	MEAN	ST.DEV <sup>n</sup> .
Men	52.5	10.36
Women	53	9.17
All	52.87	9.31

In spite of the limited number of men tested the average scores obtained by men and women differ only slightly - the difference being in favour of the women. The variability measures, the women's in particular, are a clear indication of a fairly homogeneous group. The results are superior to the George Watson's Boys' College results at 17 years and below the results for 18 years. In the Preliminary Testing, when half of the Psychology Class was tested the results were on a line with the school 18 year old average. The elimination of the best subjects here may explain the difference but the reliability of the school 18 year old results is also open to question.

#### EDINBURGH PROVINCIAL TRAINING COLLEGE.

---

All the subjects in this group were ordinary college trained non-graduate subjects, most of them falling between the age of 18 and 21 years. Only 9 men were tested and these were procured with difficulty, as a noteworthy feature of the College is the dearth of male students. The graduate men were tested in the University with the Education Class. The number of women tested affords or ought to/

to afford a good sampling of the intelligence of coming teachers of elementary school children.

It seems unnecessary to calculate separate means for men and women. The average number of marks obtained by the whole number of students tested is 48.17. The group would then fall into a position midway between the 16 year old and 17 year old boys at George Watson's Boys' College. Yet it has been stated somewhere by an American that college trained students who become elementary school teachers have as high mental ability as graduate students. Like the  $13\frac{1}{2}$  upper limit to the growth of average intelligence this may hold in America but not in Scotland!

#### HERIOT WATT COLLEGE.

The investigator carried out the test on one group of the technical students and Mr. Muir, their lecturer, applied the test at convenient periods to the other two groups. It was particularly desirable to have these groups tested as they consist entirely of male subjects who are also non-University men. In all, 59 men were tested at the Heriot/

Heriot Watt College. Their ages ranged from 16 to 24+ (see table of detailed ages).

The average score obtained by the whole group was 46.71 (st.dev.= 10.9) which is approximately equal to the 16 year old score of the Geo.Watson's Boys' College subjects. In the first test where only 17 students were tested, the average score fell between the 17 and 18 year old results from the schools. The number tested here is greater, the distribution of scores indicates no undue influence of extreme measures, so that it is just to conclude that the second result is more reliable than the first. A question may still be raised as to the possibility of this group being handicapped by a somewhat linguistic test. It is not likely. The men are at least as able to respond to this situation as well as the technical school children. One other point must be mentioned. The mean is derived from the massing together of a few cases at each of nine age groups. On this account only is the result to be read with caution.

Whether older subjects, such as extreme cases contributing to the 24+ age group, are handicapped, is another question to which GREEN would have/

have given and YERKES would give a positive answer. But the question of intelligence decline is not yet considered widely. Only "inadaptability" is considered.

# EDINBURGH UNIVERSITY EDUCATION CLASS AND EDINBURGH PROVINCIAL TRAINING COLLEGE EDUCATION CLASS

---

In the University Education Class room 188 subjects were tested. These were all graduates of some University, in the year of their training for teaching at one or other of the two Training Centres in Edinburgh. The students who are at this stage must attend an Education Class and may not attend it before graduating. St. George's Training College <sup>Sends</sup> ~~and~~ all its teachers-in-training to the University Class. The Edin. Prov. Tr. College, a much larger body, has its own Education Class for Graduation, but, since many students training there are desirous of obtaining a Diploma in Education at this period, for the procuring of which attendance at the University class is compulsory, it happens that a considerable number of them are to be found in/



in the University Class. Under a temporary arrangement these two classes were meeting together at the University. Since all were graduates, no separation of the groups has been attempted.

The table below gives the average score obtained in the test by 50 men and 138 women.

	MEAN	STANDARD DEV <sup>n</sup> .
Men	60.42	12.6
Women	58.92	10.5
All	59.30	11.2

The average of the men is slightly superior to that of the women but their variability is also greater. The 18 year old level of the school children is easily surpassed by the graduate men and women, in this as in the previous test (though the doubtful 18 year old level of the Edinburgh Ladies' College girls just passes the average of the graduate women here). On the whole this is probably as high an average level of mental ability as can be obtained by testing and large group.

COMBINED/

# SECOND INVESTIGATION.

Detailed Ages. Adults. (420 tested).

COLLEGE.		AGE.											TOTAL.
		16	17	18	19	20	21	22	23	24*			
Psychology Class	Men	.	.	.	3	2	4	1	3	1	14		
	Women	.	.	.	4	20	13	3	.	3	43		
	Men	.	.	1	3	2	2	1	.	.	9		
	Women	.	1	20	38	31	8	7	1	1	107		
Heriot Watt College	Men	2	3	4	8	9	8	9	5	11	59		
	Men	.	.	.	2	3	10	6	13	16	50		
Education Classes (Univ. & E.A.C.)	Women	.	.	.	7	8	33	52	24	14	138		
	TOTAL.	2	4	25	65	75	78	79	46	46	420		
TOTAL "	Men	2	3	5	16	16	24	17	21	28	132		
	Women	.	1	20	49	59	54	62	25	18	288		

## COMBINED ADULT RESULTS.

---

Taking all the subjects together and treating the results according to the different age groups into which they fall, it is possible to get some idea of the average growth curve of these adults. From the general table below, giving the detailed sources of the adult subjects, it can be seen that here, as in the preliminary test, there are too few men at each age for any great reliability to be attachable to the average obtained. With the women the situation is more satisfactory, at least from 19 to 22 years. When the results for men and women are taken together the numbers at 20 years and upwards are fairly satisfactory.

In the first preliminary testing, it was felt that the need for obtaining a greater number of adult subjects was great. Accordingly every available source was tapped. It was the greatest misfortune that an attempt to obtain access to groups of non-selected adult subjects met with no success. There was nothing for it but to present the results as they were found and represent the wish that other than/

AGE.	MEAN.	P.E. OF MEAN.	ST. DEV.	MEDIAN.
19	M. (53.3) W. (49.9) ALL 50.81	(1.43) (.85) .74	( 9.52) ( 9.59) 9.73	(53.83) (49.75) 51.07
20	M. (52.12) W. (50.6) ALL 50.97	(2.12) (.89) .82	(12.6) (10.29) 10.64	(49.5) (50.9) 50.77
21	M. (56.79) W. (55.98) ALL 56.23	(1.93) (1.10) .97	(14) (12.04) 12.67	(58.25) (57.43) 57.6
22	M. (51.35) W. (58.19) ALL 56.72	(2+) (.98) .95	(15.26) (11.48) 12.57	(53) (58.25) 57.53
23	M. (55.33) W. (55.8) ALL 55.58	(2+) (1.61) 1.38	(15.8) (11.97) 13.93	(57.08) (57.75) 55.8
24	M. (56.75) W. (53) ALL 55.28	(1.24) (1.84) .96	( 9.8) (11.62) 9.66	(56.5) (53.7) 55.1

The 'adult' results prior to 19 years are incorporated with the school results in the general discussion.

The curves on the opposite page show the results in another form.

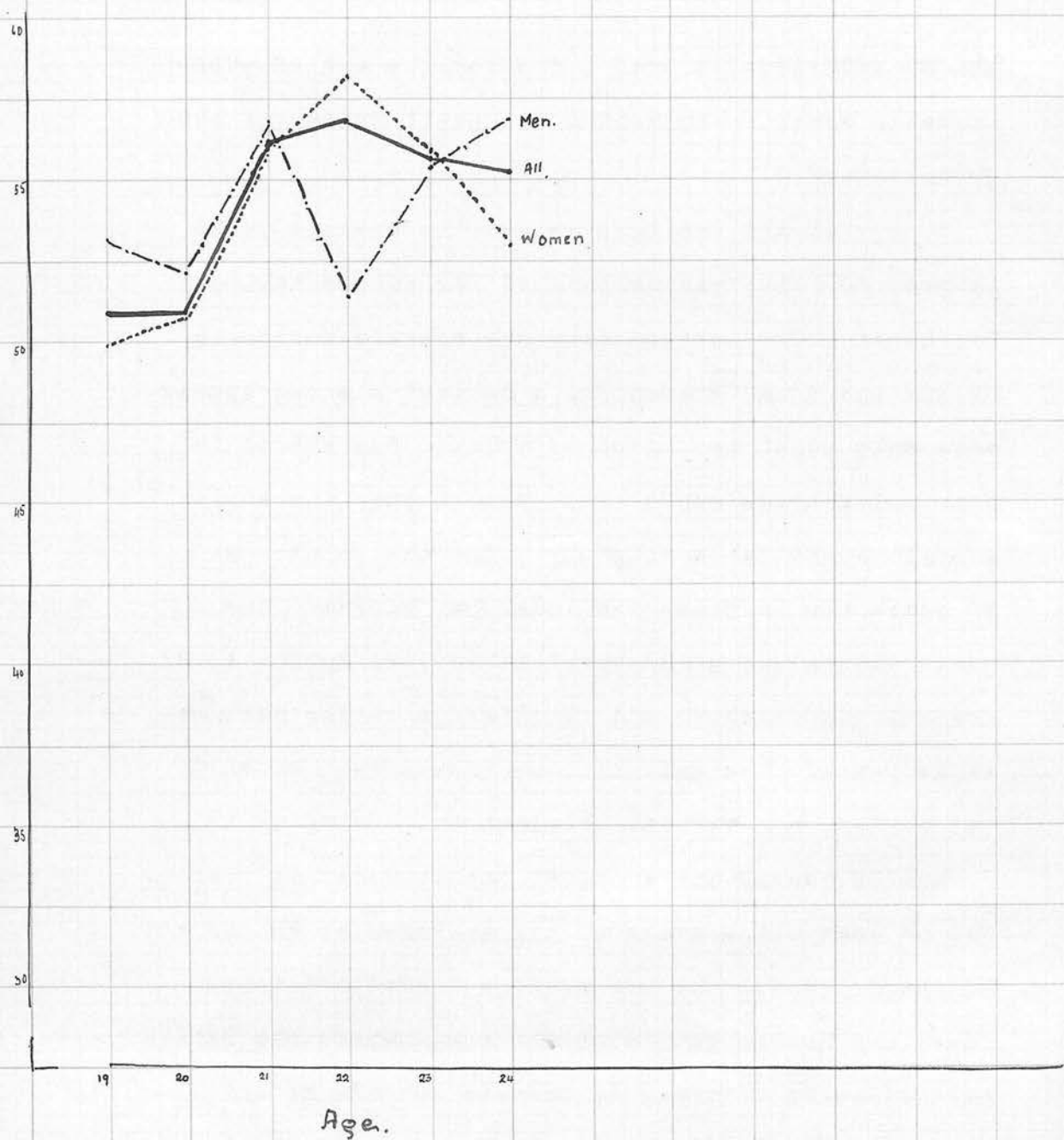
than selected subjects had contributed to the averages. The adults tested are on the average selected and therefore their average scores do not represent what would be expected of an adult group of average ability. But, if the results are not strictly comparable with the scores at the young school ages, they serve well to bring out the contrast between average and supernormal mental ability. In the table below the results obtained at each age are collected together. The numbers from which the results are obtained can be seen in the general table of distribution of ages. The adult distributions may be seen on a later page, in the discussion of combined school children and adult results. They are, on the whole, inadequate.

#### MEN.

The inadequacy of the men's distributions is seen by a glance at the page containing them. They present a flat, distended appearance. At three of the age groups the average score is the result of combining less than 20 individual scores and at the others the average score is composed of between 20 and/



# Adult Curves



and 30 individual scores. The results are of questionable worth, considering the small number of subjects tested.

At all the ages tested the average score is over 50. The variability of the groups varies. The total curve derived from the average scores is of curious form, presenting a decided zig-zag appearance when point is joined to point. The school 18 year old average was 55.24. The 19 year old adult average score falls slightly below the point. At 20 years the decrease continues; at 21 years the curve shoots up; at 22 years it is down again, at 23 years it is climbing and it climbs a little farther at 24 years. The results lend themselves better to two curves, a lower one obtained by joining up 19, 20 and 22 years, and a higher one obtained by joining the 18 year old average to the averages at 21, 23 and 24 years. Where the low average scores are found, there the Heriot Watt students are. Where the Heriot Watt students do not preponderate the age groups show a high average score. The curve exhibits the same general inclinations as the first test but the testing of more graduates at 23 and at 24 years raises/

raises the average score at these ages. The average of all the men's test results above 18 years is 54.26, showing that 18 years is approximately the average upper limit for such subjects as have been tested here.

### WOMEN

In the interpretation of women's results the sources of the subjects must be kept in mind. The scores of University and non-university students combine to give the general averages, but there are more University than Training College women. The successive age groups have not similar average scores, but the variability of the groups is remarkably steady. The lowest average score is at 19 years. At 19 and 20 years the average scores are no higher than those obtained by 17-year old Edinburgh Ladies' College girls. At these ages the results are chiefly based on the test scores of E.P.T.College women who fall into the age groups at 19 and 20 years more than anywhere else. The presence of these women at the age 19 and 20 accounts for the difference between the average scores of men and women. At 21 years the average score is increased and the curve shoots/

shoots up to precisely the same point as is reached by the men. At 22 years there is a further increase, whereas the men's curve or table of results indicates lower average scores. At 22 years, the women tested are for the most part post-graduate students. Hence their high average scores. After 22 years the averages begin to decline and the curve echoes the fact. At 23 years and at 24 years, where the curve of growth returns to where it started at 19 years, the numbers are too small for the results to have any great significance.

The general curve of women's average scores reveals the same 'trough' and 'crest' appearance as before, but with the summit of the crest at 22 years rather than at 23 years. It is based on distributions which are characterised by the fact that a smaller number of individuals contribute to them than to those at the school period. The general results, when compared with the corresponding results of the preliminary test, to show that the outcome of the second investigation is similar to the outcome of the first as far as the women's results are concerned. On the average all the adult women tested are distinctly/

distinctly above the normal.

#### COMBINED RESULTS OF MEN AND WOMEN.

---

The combination of men's and women's results at each age yields an adult general average curve. While some of the peculiarities of the single curves are smoothed away at the later ages, so that from 21 years onwards the average curve has nearly a straight line appearance, the E.P.Tr. College women's scores at 19 and 20 years are too strongly represented and the total number of men at each age is too small for the curve to show much change at the lower adult ages. Leaving out the "trough" of the curve at 19 and 20 years, there is a fair indication of a levelling up of average scores from 18 to 24 years. (The table of average scores presents the data on which the curve is based.)

At 19 and 20 years the average scores approximate to the 17 years old level of the school boys, but thereafter the tendency is all towards an approximation to the 18 year olds level. The general adult curve seems to support the finding of the preliminary test that there is a tendency for the average adult/



# NUMBER AND SOURCE OF SUBJECTS.

AGE.	MEN.	WOMEN.	MEN and WOMEN.	SOURCE.
12	107	78	185	G.W.B.C. 53 E.L.C. 13 J.C. 119
13	195	187	382	G.W.B.C. 87 E.L.C. 67 J.C. 228
14	219	174	393	G.W.B.C. 128 E.L.C. 81 J.C. 184
15	155	97	252	G.W.B.C. 129 E.L.C. 76 J.C. 47
16	117	68	185	G.W.B.C. 110 E.L.C. 64 J.C. 9 H. Watt C. 2
17	83	54	137	G.W.B.C. 80 E.L.C. 53 E.P.T.C. 1 H. Watt C. 3
18	30	41	71	G.W.B.C. 25 E.L.C. 21 E.P.T.C. 21 H. Watt C. 4
19	20	57	77	Psych. Cl. 7 E.P.T.C. 41 H. Watt C. 8 Educ. Cl. 9 E.L.C. 8 G.W.B.C. 4
20	16	59	75	Psych. Cl. 22 E.P.T.C. 33 H. Watt C. 7 Educ. Cl. 11

# NUMBER AND SOURCE OF SUBJECTS. (contd.)

AGE.	MEN.	WOMEN.	MEN and WOMEN.	SOURCE.
21	24	54	78	Psych. Cl. 17 E.P.T.C. 10 H.Watt C. 8 Educ. Cl. 43
22	17	62	79	Psych. Cl. 4 E.P.T.C. 8 H.Watt C. 9 Educ. Cl. 58
23	21	25	46	Psych. Cl. 3 E.P.T.C. 1 H.Watt C. 5 Educ. Cl. 37
24	28	18	46	Psych. Cl. 4 E.P.T.C. 1 H.Watt C. 11 Educ. Cl. 30
TOTAL	1032	974	2006.	

G.W.B.C. = George Watson's Boys College.

E.L.C. = Edinburgh Ladies' College.

J.C. = James Clark.

H.Watt C. = Heriot Watt College.

E.P.T.C. = Edinburgh Provincial Training College.

Psych. Cl. = Psychology Class.

Educ. Cl. = Education Class.

adult scores derived from more or less academic subjects to be concentrated at this level.

#### GENERAL RESULTS OF SECOND INVESTIGATION.

---

To collect together the records from all sources and present them in unified form is no mean task. In this, the second part of the investigation, all subjects may be included in the final record, since all were given the same time limit to the test of mental ability. The combined interpretation of the results will have to take many facts into consideration. The table opposite gathers together all the subjects, men and women, boys and girls, tested in the second investigation and alongside are placed the sources of the subjects. The sexes are fairly equally represented in general.

The second table gives the detailed results for all ages. The mean, P.E. mean, standard deviation and median for all ages are found there. The distributions themselves are added since they reveal the general situation clearly.

The curves of average growth for all subjects together and for men and women separately find a place near by. The higher dotted curve represents the average course of the George Watson's Boys' College/

TABLE CONTINUED

AGE	MEAN	P. E. of MEAN	STAND DEV.	MEDIAN
20	M. (52.12) W. (50.6) ALL 50.97	(2.12) (.89) .82	(12.6) (10.29) 10.64	(49.5) (50.9) 50.77
21	M. (56.79) W. (55.95) ALL 56.23	(1.93) (1.10) .97	(14) (12.04) 12.67	(58.25) (57.43) 57.6
22	M. (51.38) W. (53.19) ALL 56.72	(2+) (.98) .95	(15.26) (11.48) 12.57	(53) (58.25) 57.53
23	M. (55.33) W. (55.8) ALL 55.58	(2+) (1.61) 1.38	(15.8) (11.97) 13.93	(57.08) (54.75) 55.8
24	M. (56.75) W. (53) ALL 55.28	(1.24) (1.84) .96	(9.8) (11.62) 9.66	(56.5) (53.7) 55.1

*Sheet Reversed*

AGE	MEAN	P.E. OF MEAN	STAND. DEV.	MEDIAN.
12	M. (26.76 W. (24.64 ALL 25.87	(.65 (.83 .51	(10.01 (10.85 10.43	(25 (22.9 25.6
13	M. (28.05 W. (29.22 ALL 28.62	(.57 (.59 .41	(11.9 (12.96 11.97	(26.43 (27.16 27.08
14	M. (35.96 W. (33.12 ALL 34.70	(.60 (.75 .47	(13.16 (14.56 13.86	(34.83 (32.5 34.03
15	M. (41.8 W. (42.75 ALL 42.16	(.72 (.92 .56	(13.37 (13.51 13.3	(41.28 (39.97 42.16
16	M. (45.94 W. (51.45 ALL 47.96	(.84 (1.21 .72	(13.51 (14.9 14.14	(46.17 (53.27 48.85
17	M. (49.41 W. (54.55 ALL 51.42	(1. (1.12 .76	(13.65 (12.25 13.37	(49.86 (53.7 50.23
18	M. (54.63 W. (54.2 ALL 54.38	(1.25 (1.37 .95	(10.15 (12.95 11.9	(55.72 (52. 54.36
19	M. (53.3 W. (49.9 ALL 50.81	(1.43 (.85 .74	(9.52 (9.59 9.73	(53.83 (49.75 51.07



College and Edinburgh Ladies' College results. The lower dotted curve represents the James Clark School average course of development.

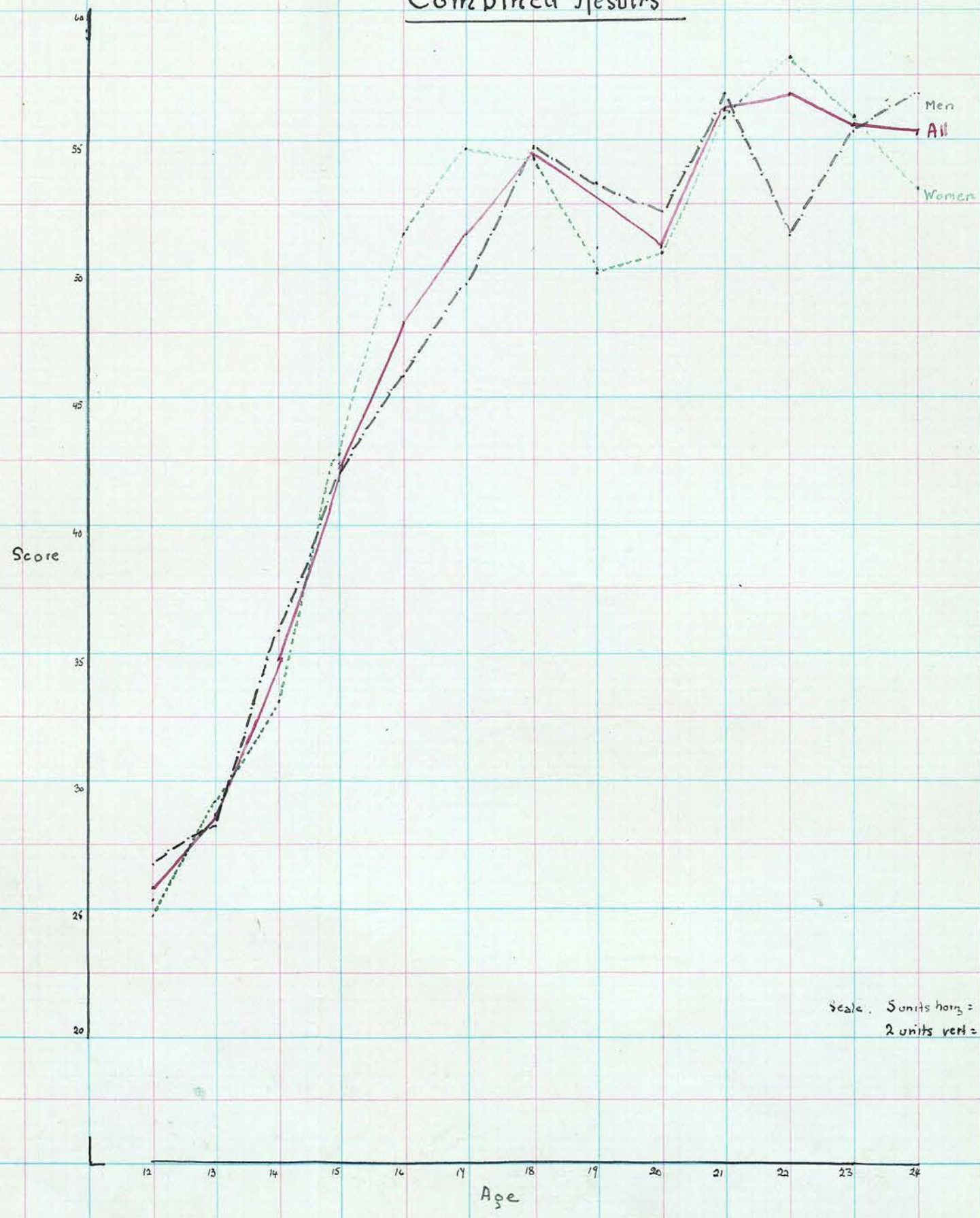
Whether it is a profitable procedure to mass together all the results remains to be seen. It is hardly justifiable unless the sources given above are kept well in memory. In all, slightly over 2000 subjects contribute to the general average and approximately the same number are men as are women. In the different age groups the representation of the sexes is not always equal as the chart shows. The boys outnumber the girls at 14, 15 and 16 years and the women outnumber the men from 18 to 23 years. Due regard of these features is necessary in the general interpretation. In the analysis that follows much repetition of previous statements must be pardoned.

#### MEN'S RESULTS.

In general appearance the curve presents a flattened form from 12 to 13 years, shows considerable and fairly even increase in average score from 13 to 15 years, continues at a slightly diminished speed up to 17 years, increases again slightly between 17 and 18 years, between 18 and 19 years drops a little lower and then presents a more or less zig-zag appearance/

2<sup>nd</sup> Test.

Combined Results



Scale: 5 units horiz = 1 year  
2 units vert = 1 mark



appearance to the end.

Contributing to the men's average score at 12 years are James Clark School boys and boys of G.W.B. College; of the former as many as of the latter. The figure (26.76) may be considered to have a fair degree of reliability. The slight increase at 13 years over the 12 years average is due to the drag upon the James Clark distribution excited by the influx of boys from another school. At 14 selection operating in the James Clark School tends to heighten the average but not exceedingly since the G.W.B.C. boys are in the majority. At 15 years chiefly owing to the preponderance of G.W.B. College subjects the average growth is nearly maintained. Up to this point the variability of the groups steadily increases as might be expected if the groups are representative. From 15 to 16 years where George Watson's Boys' College is almost solely represented the average increase in score is diminishing in spite of selection eliminating the 'poorer' subjects. Taken in conjunction with the previous increment between 14 and 15 years, which failed to exceed the 13 to 14 year increment, this is/

is evidence of a diminution in average growth which appears to have commenced after 14 years.

Despite continued selection the increase at 17 years over the 16 year old average is again less. At 18 years rigorous elimination of average subjects effects an average increase which is only intermittently surpassed in the adult years. The upper ages prove representation of purely selected subjects whose average hovers about the 18 year old level. For the determination of the upper limit of average intelligence growth they are only indirectly contributory.

The conclusion which follows from the study of averages (on the assumption that the variability of these is such as to admit of comparisons) is that average mental growth appears to be evident up to 14 years of age, diminution of growth is visible from then onwards, and the upper limit of the growth of intelligence is reached at a period assumed to be not much later than 15 years ( $14\frac{1}{2}$ - $15\frac{1}{2}$  age group). The conclusion receives support from the table below which reveals the actual differences found and the probability of their reality. The results are only extended/

extended to 17 years. Differences beyond are entirely explainable on other grounds.

LEVEL						DIFFERENCE IN AVER. SCORES	P.E.OF DIFF.
13 year aver. over 12 yr.aver.						1.29	.86
14	"	"	"	13	"	7.91	.87
15	"	"	"	14	"	5.84	.93
16	"	"	"	15	"	4.14	1.1
17	"	"	"	16	"	3.47	1.3

Up to 15 years the differences are significant being three tables, their P.E. (they are even 4 times P.E. if that figure is preferred). The average increase at 16 years over 15 years appears also to be significant, but requires to be interpreted in full remembrance that selection has entered the groups definitely at that point. There is also indication of selection at James Clark School at 15 years but the numbers are small, so that they probably do not affect the average score overmuch. Moreover, the P.E. difference can stand to be increased at 15 years without suggesting that the difference of the 15 year average over/



over the 14 year average is insignificant. The table seems to support the conclusion derived from the study of average scores and growth curves. After 16 years every selection does not result in a significant average increase in score. It seems safe to conclude that average mental growth comes to a stop soon after 15 years. Further slight confirmation lies in the fact that Heriot Watt College men (59 in number) reached on the average only the 16 year level and they are presumably at least slightly above the average in mental ability.

The table of percentages differences between the groups below requires interpretation. It is useless to extend it beyond 18 years.

---

Percentage of 13 year olds reaching or surpassing the median of the 12 year olds	53
Percentage of 14 year olds reaching or surpassing the median of the 13 year olds	73
Percentage of 15 year olds reaching or surpassing the median of the 14 year olds	62
Percentage of 16 year olds reaching or surpassing the median of the 15 year olds	60
Percentage of 17 year olds reaching or surpassing the median of the 16 year olds	57
Percentage of 18 year olds reaching or surpassing the median of the 17 year olds	66

---

Apart/

Apart from the figure at 12 years which requires no comment, there is a gradual decline of percentage increase up to 17 years. But the figure at 16 years (60%) shows no appreciable decrease over the former percentage on account of the counteracting influence of selection. Where more rigorous selection enters at 17 years only a slight decrease in percentage increase is shown. Consequently it cannot be assumed that the figures are to be literally interpreted as indicating the true general movement of the successive age groups. The approximation to equality should come considerably before 17 years. That it would come at approximately 16 years is quite in keeping with the earlier trend of the curve.

The conclusions, rough as they are, are the same by whatever method they are derived.

### WOMEN'S RESULTS.

---

The average curve of women's results placed besides the men's curve, exhibits in general, the same form as the men's. The interval between twelve and thirteen years is larger than in the boys' curve, owing to the larger number of James Clark School pupils contributing to it. A slightly smaller increment in average growth is apparent between 13 and 14 years; between 14 and 15 years the ascent of the curve is steeper and between 15 and 16 years the increment is again slightly decreased. From 16 years to 18 years only slight increase is observed, and at 19 and 20 years a marked descent in the curve is visible. From 20 to 22 years the curve is again on the upward path, but after 22 years it is descending to 24 years.

Turning now to the actual figures on which the average growth curve is based, it may be seen that the average score at 12 years is smaller than the average score of the boys at the same age.  
by/

by reason of the number of James Clark School pupils represented there. Only 13 Edinburgh Ladies' College girls were tested. Between 12 and 13 years the indications are that average mental growth is still apparent. In Edinburgh Ladies' College an extreme individual heightens the average at 13 years, and in James Clark School 'poor' subjects diminish the average. The two facts may to some extent, counteract each other. At 14 years a further increase in average score is still indicative of mental growth, but the increase is small so that apparently the 'slowing down' period of mental growth is approached here. Up to 14 years the reliability of the average measures is established. The variability measures exhibit the usual gradual increase from age to age. At 13 the larger size of the girls' standard deviation is due to the extraordinary superior subject. At 14 years the variability, still more different from that of the boys is probably largely accountable to the different number of subjects tested in each sex. After 14 years, not only does selection become apparent and influence all scores, but the Edin. Ladies' College score peculiarities exert an unfortunate influence on the average/

average score. Edin. L. College subjects contribute almost entirely to the general averages of the women's groups up to 18 years. At 15 years, the average score, if attributable to average mental growth, must also be due, in part to selective influences operating at least in the James Clark School. They may also affect the Edinburgh Ladies' College average score, to some slight extent. The size of the interval between 14 and 15 years still appears large even when selective influences account for some of its size. It is not improbable that the reason lies with the influx of pupils which effects a lowering of the 14 year old average and, as a consequence, effects a highering of the 15 year old average. Again, in considering the average score at 16 years due allowance must be made not only for selective influences, but for the increased average of 16 year old Edinburgh Ladies' College pupils due to the omission of one group of 'poor' children. These two points noticed, the average at 16 years cannot be literally interpreted. The true course of average mental growth is in fact obscured after 14 years. In spite of severe selection at 17 years the average increase at that point/



353

point is only 3.1 points, that is no greater than the average increment between 13 and 14 years. In all probability none of the increase is due to actual mental growth. At 18 years no increment exists, at 19 years there is a positive diminution, and owing to the number of Moray House women tested, at 20 years the same average is found as at 19 years. The heightened average scores beyond 20 years are due to the selected nature of the subject tested. In the adult years the average growth curve manifests first a 'trough' and then a 'crest' appearance. The averages at 18, 23 and 24 years as the P. E. Means show are unreliable on account of the small number of subjects tested.

As the actual averages go, and with consideration of the facts that require to be noticed at the different levels or ages, the findings elsewhere expressed, seem to be supported and mental growth, slight after 13 years, does not appear to be continued long after 15 years.

The table of actual differences below support this view. The table need not be extended beyond 17 years.

Table/

LEVEL							DIFFERENCES	P.E. of DIFF.
13 year olds over 12 year olds							4.58	1.01
14	"	"	"	13	"	"	3.9	.99
15	"	"	"	14	"	"	9.63	1.19
16	"	"	"	15	"	"	8.70	1.51
17	"	"	"	16	"	"	3.10	1.62

Significant differences are apparent at 13 and 14 years. At 15 years the difference over 14 years, is as previously explained, probably increased by reason of a lowered average at 14 years. At 16 years the facts, now so often reiterated regarding the Edinburgh Ladies' College results, explain the size of the difference. The P.E. of difference is scarcely valid here by itself but requires this additional explanation. The situation at 17 years (P.E. 1.62) confirms the suggestion made previously that the difference is not due to mental growth but to the operation of selection. The extension of the table is unnecessary since the adult subjects are all selected. The distributions of adult scores emphasise the 'trough' and 'crest' appearance of the adult curves. The/

The 19 and 20 distributions are distinctly lower in general form than those above these ages. The distributions, showing the general advance from age to age point to an increase in mental growth up to 17 years. This as has been seen, cannot be literally interpreted as selection obscures the course of real growth above 14 years.

As a last avenue of approach, the percentages of each age reaching or exceeding the Median of the age below are presented. But, these exhibit the same weaknesses as the other tables.

PERCENTAGE of										YEAR OLDS
13 yr. olds reaching or surpassing										
the median of the 12										61.2
14	"	"	"	"	"	"	"	"	13	60
15	"	"	"	"	"	"	"	"	14	79
16	"	"	"	"	"	"	"	"	15	73.5
17	"	"	"	"	"	"	"	"	16	44
(and so on.)										

The first two percentages are satisfactory pointing to a continuance of growth. The first percentage is lowered by reason of the aforementioned influx/

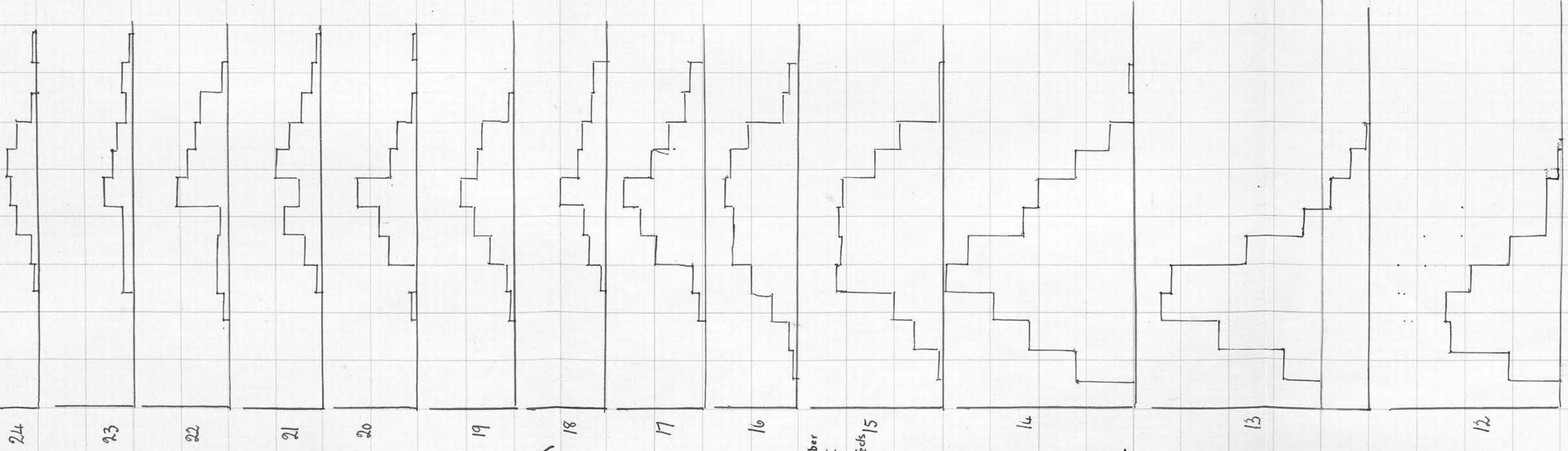
influx of James Clark School Girls at 13 years. The second percentage is similarly heightened. The percentage increase at 15 years over 14 years at once shows itself to be peculiar. It should rather be lower than the preceding percentage. Again the exceeding size of the increase is traceable to the lowered 14 year old average which makes the difference larger than it should be. It is also due, in part, to the forces of selection at the James Clark School and perhaps also at Edinburgh Ladies' College. The percentage at 16 years is probably not unduly large since the elimination of subjects of lower intelligence is marked in Edinburgh Ladies' College. What is of interest is the 17 year old percentage (44%) which first indicates a group exhibiting no general improvement over the preceding group. After 17 years at least, as distributions or percentage increases show, average growth is non-existent. But, there are reasons for lowering the figure to 16 or even 15 years. The 14 year old percentage is 60 and is probably over-large. The 17 year old percentage is 44. The latter percentage occurs where selection is rigorously operative. It is, thus, more reasonable to suppose that/



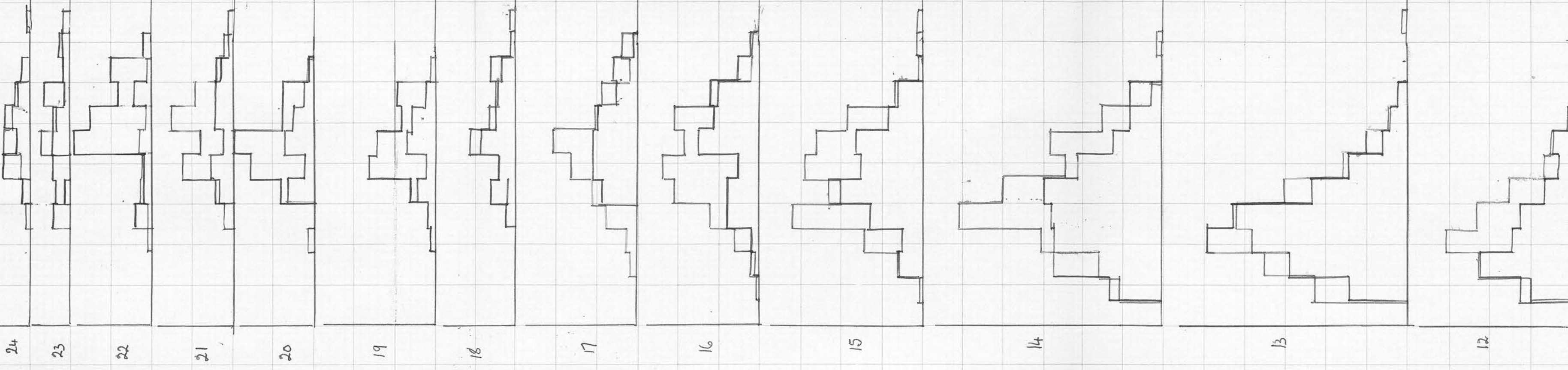
2nd Test

Final Results

All.



Men & Women.



Scale 3 units horiz. = 1 class interval (7 marks)  
5 units vert. = 10 subjects

Red - Boys.  
Green - Girls



that average mental growth ceased to grow at a point nearer 14 years than 17 years. Further, the table of differences indicated only a small improvement from 13 to 14 years. By 15 years that may become negligible, or, if not at 15 years, then proximate to 15 years. These two considerations establish a conclusion which is more favourable to a pre 16 year old cessation than to a post 16 year old cessation. Approximately 15 years, and not below 15 years is the safest place to suggest as the level.

#### ALL SUBJECTS.

---

The average growth curve resulting from the combination of 974 women's test scores at the different age groups from 12 to 24+ and of 1032 men's test scores correspondingly interspersed is drawn beside the separate men's and women's curves. Generally speaking, it exhibits the same characteristics as the curve for each sex. From 12 to 13 years it presents a more flattened appearance than afterwards. From 13 to 16 years average mental growth appears to be progressing. From 16 to 18 years there is a slight slowing down, at 19 and 20 years a marked descent seems to interrupt the true course of the curve, but from 21 years onwards the old line is re-established and the curve assumes a flattened appearance at approximately the 18 year old level/

level. As stated previously, the appearance of the general average growth curve is deceptive, suggesting at first sight an upper limit for the growth of intelligence at 18 years or thereabouts, whereas a knowledge of the real facts is disillusioning. It is the same knowledge of the facts which suggests that the combination of all the test scores under each age group conceals the most interesting feature of average mental growth that the study reveals at the earlier years. As the result of adding on the James Clark School results to the George Watsons' Boys' College and Edinburgh Ladies' College results, allowance for the influence of selection has to be made sooner than when the technical school results are omitted. (eg. in the Preliminary (Final) results) Yet the indications are that the averages obtained are reliable except at 23 years (P.E. 1.38). The preponderance of James Clark School scores at each of the earlier years makes the curve follow a parallel course to that exhibited by the James Clark School average curve (lower dotted curve) whereas the Edinburgh Ladies' College and G.W.B College combined curve (higher dotted curve) presents no such appearance of rapid, steady development. The truth/

truth is that in the final results the question of selection enters, to some extent or other, nearly all along the lines till, at the adult years the experiment becomes an indication of superior intelligence worth, without a doubt. Consequently a description of the state of affairs at each point, tedious in itself, becomes even worse so on account of the different facts that must be taken into consideration. The onus of the full description, moreover, lies with the experimenter, whose interpretation, it must be assumed, is the truest interpretation of the facts.

As far as possible, the attempt has been made to get at the truth about the factors affecting the results, by questioning the teachers and authorities, with whom, at various times during the periods of testing, the experimenter came in contact. Understanding teachers sometimes supplied invaluable information. In one way or another the supplementary information necessary was obtained. Nowhere is the additional information more necessary than in the consideration of the trend of the general curve. Incidentally, the general curve and the two separate curves, for boys and girls, lie nearly on top of one another/

another up to 15 years but after that age the general average curve falls between two curves that tend at times to diverge and at times to curve together.

From 12 to 13 years there is an average increment which is very roughly only three points. In some small measure the selected Edinburgh Ladies' College subjects of 12 years of age, tends to lessen the size of the interval as also does the preponderance of 'poor' James Clark subjects at 13 years. From 13 to 14 years the average increment is greater but the lowered 13 year old score and the possible selection of James Clark pupils at 14 years, this time serve to augment the interval. Presumably, during this period from 12 to 14 years average growth is at least noticeable in some degree. The variability measures of the groups increase gradually to 15 years, but are usually large on account of the large Edinburgh Ladies' College variability. Still, the comparison of groups at all stages is legitimate from this point of view. Between 14 and 15 years the slightly larger average increase in score is due to the entrance of selection at the James Clark School and also, probably, at Edinburgh Ladies' College. The true average increment would probably be smaller than at the previous ages. After 15 years the average increase becomes gradually less, but the diminution in/  
in/



in average growth probably appears less than it really is because of selective forces applying alternately now to Edinburgh Ladies' College and G.W.B. College as well as to James Clark School with increasing rigour. For this reason the average increments cannot be taken at their face value. Rather, the apparent decrease in average increment from year to year must be reckoned as greater than the real average increment from year to year; that is to say, as less than 5.8 units (see following table) between 15 and years, as less than 3.46 units between 16 and 17 years and as less than 2.96 units between 17 and 18 years. But even with this recognition the situation is not revealed, and the root of the trouble lies in the unreliability after 16 years, of the girls' results which go to produce over large average increments at and beyond this age.

The growth then from 16 years upwards is of doubtful validity, as all the facts go to prove, and its size is probably less than the figures indicate. By 17 years there is no doubt that average mental growth has ceased since the average scores at 19 and 20 years, obtained chiefly from selected subjects at a University or College only reach the 17 year old level. Since the measures at 16 and 17 years are over size, it is difficult to locate the average upper limit/



limit of mental growth with exactitude from the combined results. A slowing down after 14 years can be observed, which would probably enter sooner, did not the preponderance of James Clark 'poor' boys and girls weigh down the 13 year old average. A considerable amount of the average difference between 15 and 16 years must be attributed to the factors of selection, and when these are allowed for, and the unusual influence exerted on the 16 year old score by the Edinburgh Ladies' College average, it will probably be impossible to find more than a trace of average mental growth.

The analysis is supported by the table of differences for all ages, outlined below.

LEVEL		DIFFERENCE		P.E.	DIFFERENCE
yr olds		yr olds	-		-
13	over	12	2.75		.20
14	"	13	6.08		.19
15	"	14	7.46		.23
16	"	15	5.80		.28
17	"	16	3.46		1.04
18	"	17	2.96		1.21
19	"	18	- 3.57		1.2
20	"	19	.16		1.1
21	"	20	5.26		1.2
22	"	21	.50		1.35
23	"	22	- 1.15		1.67
24	"	23	- .30		1.67

The only significant differences are to be found at the early places up to 16 years. (The difference at 17 years is of doubtful validity) and between 20 and 21 years. This is what might be expected from a knowledge of all factors contributing to the final results. Up to 14 years real differences may/

may be expected and probably also between 14 and 15 years though the difference here indicated is over large. At 16 years the truth is hidden. The difference indicated is not a real difference as it stands. Whenever the trouble with the Edinburgh Ladies' College results entries, there the reliability is at question. The difference between 20 and 21 years is probably of interest though the students are selected. It is due to the abrupt change in the type of students tested. They are Edinburgh Provincial Training College women at 20 years and chiefly graduate and post graduate subjects at 21 years. The table serves in a quantitative fashion to corroborate fully the facts established by the analysis of the curves with the aid of additional information. It appears that there is a real difference from age to age up to 15 years and that it gradually falls off after 16 years. Better results would probably indicate slight or zero growth soon after 15 years of age.

In order to express otherwise the general movement of the age groups the table of percentage increase may be added. It reveals the same weaknesses as the rest of the results.

Table/

TABLE

---

Percentage of 13 yr olds reaching or surpassing the median of the 12 yr olds	54.4
Percentage of 14 yr olds reaching or surpassing the median of the 13 yr olds	65.6
Percentage of 15 yr olds reaching or surpassing the median of the 14 yr olds	65
Percentage of 16 yr olds reaching or surpassing the median of the 15 yr olds	62.1
Percentage of 17 yr olds reaching or surpassing the median of the 16 yr olds	56.2
Percentage of 18 yr olds reaching or surpassing the median of the 17 yr olds	60

---

It may be assumed that the figures at the first two age groups 13 and 14 years are satisfactory. Then, it can be seen that at the third age group the percentage (65) is too high, probably again due to selection chiefly at James Clark School. Similarly at 16 years the percentage (62.1) is doubtless over large, though in comparison with the previous percentage it appears more satisfactory. Towards 17 years the groups are becoming more generally equal and at 18 years the increased percentage/

percentage is not unexpected since selection is severe at this point. The first indication of no change in the general group movement, occurring after 17 years must be noted. Taking account of the selective factors entering after 14 years, it is not unlikely that the first real indications of average cessation of mental growth occur at least a year earlier and it is not unlikely that real average growth is little apparent beyond 15 years (i.e.  $14\frac{1}{2}$  to  $15\frac{1}{2}$  age group).

#### COMPARISON OF TEST I WITH TEST II.

---

No claim of perfection is made on behalf of either test. The construction is arbitrary, and the scoring largely so, but it is believed that the test elements are only on the whole suitable for the mental testing of adult subjects. The difficulty of devising new and good tests for adult mentality is keenly felt and it seems to the writer that there is a real tendency for group tests to sacrifice real quality of performance to speed. For this reason the tests here employed were/



were graded and, even apart from the time limit, which is so attached that none shall finish, it is probable that some of the material would baffle the most keen. The code test, for example, demands concentration and concentration is at least closely related to intelligence. The later analogies are stiff also. But still, as regards the tests generally employed with adults, it may be that Roback's (1) criticism of their tendency to ignore the penetrative qualities of the mind, is true. In a time limit test the danger of superficiality may be greatest.

As far as the total results go the first and second tests seem to corroborate each other's findings. The second test, slightly longer, seems also to spread the results more than the first but the difference is slight. It had been intended to equalise the tests by some rough method, but, taking all inequalities into consideration, it seemed, to be an unprofitable procedure. The general graphs may suffice to show that the tests at least are non-contradictory in their findings. The same parallelism/

---

(1) ROBACK: Journal of Education Psychology Nov. 1921.

parallelism of average mental growth curves for each sex characterises both up to 16 years and tends to continue to 18 years. The adult results exhibit the same general course, though inadequate numbers tested and lack of homogeneity in the material used make particular inequalities apparent at certain ages. By means of correlation methods the self correlation of the tests or the correlation of Test I with Test II was obtained for each main group of children tested. The correlation was .82 with George Watson's Boys' College, .78 with Edinburgh Ladies' College and .76 for James Clark School. Again these results are fairly satisfactory considering the imperfection of the tests, and observing also that some nine months intervened between the application of the two tests.

#### CORRELATION WITH OTHER TESTS.

---

It happened that one Class of Edinburgh Ladies' College subjects received NATIONAL INTELLIGENCE Test at some time previous to the application of the second test. Accordingly a correlation was worked out between the results of the standardised National Intelligence Test and the second test employed/

employed here. The correlation obtained was .76. The numbers were unfortunately small, being only a little over 30, but the figures are satisfactory for this group. Similarly, one class of George Watson's Boys College pupils last received the OTIS Group test. The correlation obtained was .75. Again the result is satisfactory, though the small numbers must be noted.

#### INTELLIGENCE SCORES AND ESTIMATES OF INTELLIGENCE.

---

Only with two adult sections were correlations found between test scores and other estimates of intelligence. At Edinburgh Provincial Training College the total marks of each student tested in the ordinary Two Year Course were put against their test scores. About 80 contributed to the correlation of .52.

In connection with the Heriot Watt subjects Mr Muir, their Lecturer, before seeing the test results, estimated, as near as possible, the intelligence of the 59 subjects tested. The correlation between his estimate and the test score was .49, which/

which is as high as is usually obtained, from adult subjects. Moreover Mr Muir complained that he had been set "an exceedingly difficult task" and that his estimate was probably of limited worth on account of it being based on a somewhat specialised contact with the students in the engineering classes.

#### SEX DIFFERENCES.

---

Thus, as in the previous test it must be assumed that the real sex difference discovered is very slight. Previously direct comparison of George Watson's Boys College with Edinburgh Ladies' College was not effected but in connection the comparison may be made. In the tables below the means of the boys and girls at each age up to 18 years are entered and the difference between boys and girls scores with P.E.s appended

AGE	MEAN	DIFFERENCE	P.E. of DIFFERENCE
12	Boys 31.60 Girls 39.53	7.93	
13	Boys 35.62 Girls 37.32	1.7	1.3
14	Boys 40.31 Girls 40.38	.07	1.22
15	Boys 43 Girls 44.15	1.15	1.3
16	Boys 46.89 Girls 52	5.11	1.4
17	Boys 50.11 Girls 54.18	4.07	1.6
18	Boys 55.24 Girls 59	3.78	2.3



AVERAGE SINGLE TEST SCORES. (Contd.)

		①	②	③	4	⑤	⑥	⑦	⑧	9	⑩	⑪	⑫	⑬	14	⑮	⑯	⑰	⑱	20	Code
18 (Cont.) Men		5.6	3.2	6	2.4	3.1	4.4	4	3.8	6	3.6	4	3.2	3.4	3.1	4	.4	-	-	.6	.8
Women		5.5	2.5	5.4	3	3.3	5.25	2.8	3.5	7.1	3.9	3.7	3	2.15	1.8	3.3	.8	1	.5	2.6	.8
19 Men		5.8	2.2	5.8	2	3.8	5.1	2.6	2.7	7.4	4.9	4.1	1.8	1.8	2	2.3	.9	1.2	.4	2	.5
Women		5.9	2.6	4.9	2.2	3.9	5.1	2.9	3.4	7.3	4.6	4.2	2.1	2.2	1.3	2.8	.5	.1	.5	1.4	.7
20 Men		5	2.7	5.8	3.7	4.3	5.5	3.3	3.2	7	4.9	3.9	1.7	1.5	2	3.5	.5	.5	.7	2.6	.4
Women		5.7	2.3	5.7	2.2	4.3	5.2	2.7	3.8	8	4.8	3.8	2.3	2.2	1.6	2.7	.5	.2	.2	1.3	.2
21 Men		5.8	2.2	5.6	3	4.5	5.2	3	4.5	8	4.1	4.2	2.4	3.1	2.2	3.1	.9	.1	.1	.8	-
Women		6	2.5	6	1.8	4.6	5.3	3.2	3.9	9.5	5	4.3	3.2	2.5	2.2	3.3	.9	.6	.6	1.6	.6
22 Men		5.7	2.5	6	2.8	3.6	5.1	3	4.2	7	4	3.5	1.8	3	2.3	2.7	.6	.9	.3	.8	.9
Women		5.8	2.6	5.9	2.4	4.7	5.4	3.4	3.6	9.3	5	4.9	3	3	2.1	3.2	1	.7	.5	1.1	.3
23 Men		5.5	2.6	5.6	3	3.9	5.5	3	3.9	8.3	4.2	4.3	2.5	2.5	2.5	3.1	1.6	.7	.9	1.7	.19
Women		5.5	2.6	5.7	1.9	4.1	5.1	3	4.2	9.3	4.5	4.4	3.1	3.1	1.9	2.6	.5	.3	.3	.8	.2
24 Men		5.8	2.8	5.9	2.6	4.1	5.1	3.7	4.4	8.9	4.3	4.2	3.7	2.2	2.4	3	.9	.3	.7	1.1	.6
Women		5.9	2.2	4.8	1.7	4.1	5	3.2	3.2	9.2	5	4.4	3.5	2.4	2.5	3.2	1	-	-	1	-

The table shows that differences are insignificant up to 15 years. At 16 years the difference is a false representation of facts and beyond year 16 no appreciable differences are apparent. On the whole these results go to show that sex differences in reality and with reference to the total score are markedly slight.

Curiosity and a persistent doubt regarding the higher Edinburgh Ladies' College scores at all ages, led to an analysis of the test elements themselves. The task was not unprofitable, though very long.

On the opposite page the results of the analysis are reposed. The figures 1, 2, 3, 4, etc. refer to the tests in their natural order. One is analogies, two directions, and so on. The coloured circles and marks round the figures, indicate tests of the same nature. For example, 1, 6, 11, and 16 are analogies. The figures within the spaces indicate the average marks obtained in each test. The details of the chart need not be entered into. A study of the averages indicates that George Watson's Boys' College, and the Edinburgh Ladies' College pupils obtain approximately the same average score at each/

each test all along the line. This is true generally, and also that the James Clark School children tend to make lower average scores. The majority of the tests show an increase in average scores up to about 20 years. This is an indication of their satisfactoriness for this problem. Directions and analogies tests seem to be especially good. There is also a slight tendency, particularly in the James Clark School results for boys to be superior to girls in number of completion questions and girls to be superior to boys in the proverbs test. But undoubtedly the main interest in the sheet is to be found in the study of test question 9. Here lies the secret of the girls superiority. In this opposites test they have averages which clearly overtop the boys averages. The differences are noticeable all through the school ages. Two suppositions follow. Either the girls are accustomed to work of this nature, or the test favours " a linguistic sex." Probably the latter is the truer supposition since in the analogies tests their superiority, though not marked, is often apparent. Taken in conjunction with the slight superiority of James Clark School girls in the linguistic test this seems to prove that the test has a linguistic bias. Analysis might reveal the same fact in the first test. This however, /

however, should not invalidate the test conclusions. All the subjects tested are on fairly equal ground and they are beyond the stage of childhood proper.

#### SOCIAL STATUS.

The general curve, which is the resultant of all scores obscures the differences which were seen in the first test. But, regarding the curves obtained from James Clark School on the one hand and George Watson's Boys College and Edinburgh Ladies College combined on the other, it is possible to observe once again the influence of social status on the scores. That the 15 year old, James Clark scores are not approximately equal to the others, is probably due in part to the difference in test elements in the second test. There is good reason to believe that the influx of 'poor' subjects accounts for the remainder of the difference. The tendency for elimination to decrease the difference and bring the curves closer together as the years passed is still apparent. The probability is that the trend of James Clark curve is probably a misrepresentation of real average growth at the adolescent years,

GENERAL/



## GENERAL CONCLUSIONS.

---

It is reiterated that the study of averages makes the assumption that each age group represents what the preceding age group will become and, further that a study of this nature reveals directly, nothing regarding the individual difference in rate of growth or individual variations of all kinds that may arise, except in exceptional cases. The limitations thus imposed upon the study of mental growth changes are, therefore obvious. But for the study of the main problem here, the method of group testing is not entirely without justification. Neither is it without value. So much emphasis is laid on the imperfections of the method that it is forgotten how few good tests are available even for the individual testing of older subjects. What it is desired to know here is the general movement of the average curve from age to age. It cannot be claimed that the choice of subjects here employed facilitated the discovery of the upper limit of the growth of intelligence but at least with additional true information, some idea of the age of mental arrest for average subjects has been deduced.

The conclusions from the study of  
GEORGE/



GEORGE WATSON'S BOYS' COLLEGE results are that a slowing down of average mental growth is apparent after 14 years and that since selection operates from 15 years onwards, there is no sign of a counter balance in increased average score from then onwards. (This result applies to both tests.) The assumption then is, that after 15 years the slowing down will continue and absolute zero average growth be apparent by 16 years. The same conclusions can be derived from the results of the first test applied to EDINBURGH LADIES' COLLEGE pupils, and under more favourable circumstances would have been directly derivable from the Second test.

The results obtained from both tests applied to JAMES CLARK SCHOOL indicate a slightly lower point of arrest but the conclusions are handicapped by selective forces operating at both ends of the school so that it is difficult to be definite in fixating the age of mental arrest. Presumably shortly after 15 years is the best point to suggest.

When due regard of the superior social status of the GEORGE WATSON'S BOYS' COLLEGE and EDINBURGH LADIES' COLLEGE subjects is taken and also of the average social status of the JAMES CLARK subject/

subject it is inevitable that adherence to 15 years rather than 16 years should follow. The conclusion that the upper limit of the growth of average intelligence is reached at or about 15 years receives particular confirmation from the fact that DUNFERMLINE PHYSICAL TRAINING COLLEGE subjects in the first year of testing did not reach and HERIOT WATT COLLEGE subjects in the second year of testing did not exceed the 16 year old level. Of all adult subjects they most nearly approach average mentality. The adult results in conclusion with EDINBURGH PROVINCIAL TRAINING COLLEGE indicate only a 17 year old level for these subjects and they are undoubtedly of superior intelligence on the average.

Regarding the tendency of the general curve to bend noticeably at 18 years and the tendency of the adult average curve to approximate to the 18 year old level only a few words need be said. The fact that the subjects are generally of superior intelligence beyond 18 years seem to indicate that the bend of the curve is not lacking in significance and that it is a suggestion of an average upper limit of the growth of superior intelligence occurring at 18 years. Perhaps the desirability of finding such a result is a/

a bias. The curve form may only be due to chance; but it is thought not. The 18 year old school children are rigorously selected and also the post graduate subjects. The addition of lower subjects at the adult years is slight. Undoubtedly the chief reason why the suggestion cannot be pushed farther is, that an inadequate number of adult subjects have been suggested.

The outcome of the investigation is a refutation of the finding of the American Army testing, that on the average mental growth ceases after  $13\frac{1}{2}$  years. The indications of these experiments are that on the average, mental growth continues at a very slow rate from 12 to 15 years, and that after 15 years no appreciable improvement can be observed. The conclusions are not contradictory to the findings or opinions of the best investigators in the realm of mental testing.

many children at the JAMES CLARK school come there only to occupy their last two years of compulsory school attendance is some proof that they are not selected. That which selects them, and which selects always in such cases, is parental goodwill, good sense, or the material possibility of transfer to such a school. Individual cases differ but the general rule remains. The value of technical education has hardly yet made its appeal to the home environment except in a desultory fashion. This is not a denunciation of the school. Far from denunciation it merits the most praise that can be given it. The writer firmly believes in its worth. It receives children from the elementary schools who are strangely dependent on others. Independence they have not. The curious thing is that this school teaches them what they lacked. The difference at the early stages of the school and at the later was so apparent to the writer that it called for comment. The answer was that the children were taught "to think for themselves". The exact bearing of the by-talk on the question at issue may be wondered. It has been indulged in to point out that some part of the ability/